

Advanced Classical Mechanics

Prof. Iain W. Stewart

Classical Mechanics III
MIT 8.09 & 8.309

Editors: Francisco Machado, Thanawuth Thanathibodee,
Prashanth S. Venkataram

©2016 Iain W. Stewart

Last Compiled: September 23, 2021 – 11:48

Contents

1	A Review of Analytical Mechanics	1
1.1	Introduction	1
1.2	Lagrangian & Hamiltonian Mechanics	1
1.3	Symmetry and Conservation Laws	12
1.4	Constraints and Friction Forces	16
1.5	Calculus of Variations & Lagrange Multipliers	22
2	Rigid Body Dynamics	30
2.1	Coordinates of a Rigid Body	30
2.2	Time Evolution with Rotating Coordinates	34
2.3	Kinetic Energy, Angular Momentum, and the Moment of Inertia Tensor for Rigid Bodies	38
2.4	Euler Equations	45
2.5	Symmetric Top with One Point Fixed	47
3	Vibrations & Oscillations	53
3.1	Simultaneous Diagonalization of \hat{T} and \hat{V}	54
3.2	Vibrations and Oscillations with Normal Coordinates	56
4	Canonical Transformations, Hamilton-Jacobi Equations, and Action-Angle Variables	61
4.1	Generating Functions for Canonical Transformations	61
4.2	Poisson Brackets and the Symplectic Condition	67
4.3	Equations of Motion & Conservation Theorems	70
4.4	Hamilton-Jacobi Equation	74
4.5	Kepler Problem	80
4.6	Action-Angle Variables	84
5	Perturbation Theory	93
5.1	Time Dependent Perturbation Theory for the Hamilton-Jacobi Equations	94
5.2	Periodic and Secular Perturbations to Finite Angle Pendulum	96

CONTENTS

5.3	Perihelion Precession from Perturbing a Kepler Orbit	99
6	Fluid Mechanics	105
6.1	Transitioning from Discrete Particles to the Continuum	105
6.2	Fluid Equations of Motion	110
6.2.1	Continuity Equations	110
6.2.2	Ideal Fluid: Euler's Equation and Entropy Conservation	113
6.2.3	Conservation of Momentum and Energy	114
6.3	Static Fluids & Steady Flows	116
6.4	Potential Flow	122
6.5	Sound Waves	126
6.6	Viscous Fluid Equations	129
6.7	Viscous Flows in Pipes and Reynolds Number	132
6.8	Viscous Flow Past a Sphere (Stokes Flow)	136
7	Chaos and Non-Linear Dynamics	143
7.1	Introduction to Chaos	144
7.1.1	Evolution of the system by first order differential equations	144
7.1.2	Evolution of Phase Space	146
7.1.3	Fixed Points	147
7.1.4	Picturing Trajectories in Phase Space	150
7.2	Bifurcations	154
7.2.1	Saddle-Node Bifurcation	155
7.2.2	Transcritical Bifurcation	158
7.2.3	Supercritical Pitchfork Bifurcation	160
7.2.4	Subcritical pitchfork bifurcation	161
7.3	Fixed Points in Two-Dimensional Systems	165
7.3.1	Motion Near a Fixed Point	165
7.3.2	Systems with a conserved $E(\vec{x})$	172
7.4	Limit Cycles and Bifurcations	174
7.4.1	Poincaré-Bendixson Theorem	176
7.4.2	Fixed Point Bifurcations Revisited and Hopf Bifurcations	177
7.5	Chaos in Maps	183
7.6	Chaos in Differential Equations, Strange Attractors, and Fractals	190
7.6.1	The Lorenz Equations	190
7.6.2	Fractals and the Connection to Lyapunov Exponents	192
7.6.3	Chaos in Fluids	195

Chapter 1

A Review of Analytical Mechanics

1.1 Introduction

These lecture notes cover the third course in Classical Mechanics, taught at MIT by Professor Stewart to advanced undergraduates (course 8.09) as well as to graduate students (course 8.309). In the prerequisite classical mechanics II course the students are taught both Lagrangian and Hamiltonian dynamics, including Kepler bound motion and central force scattering, and the basic ideas of canonical transformations. This course briefly reviews the needed concepts, but assumes some familiarity with these ideas. References used for this course include

- Goldstein, Poole & Safko, *Classical Mechanics*, 3rd edition.
- Landau and Lifshitz vol.6, *Fluid Mechanics*. Symon, *Mechanics* for reading material on non-viscous fluids.
- Strogatz, *Nonlinear Dynamics and Chaos*.
- Review: Landau & Lifshitz vol.1, *Mechanics*. (Typically used for the prerequisite Classical Mechanics II course and hence useful here for review)

1.2 Lagrangian & Hamiltonian Mechanics

Newtonian Mechanics

In Newtonian mechanics, the dynamics of a system of N particles are determined by solving for their coordinate trajectories as a function of time. This can be done through the usual vector spatial coordinates $\mathbf{r}_i(t)$ for $i \in \{1, \dots, N\}$, or with generalized coordinates $q_i(t)$ for $i \in \{1, \dots, 3N\}$ in 3-dimensional space. These generalized coordinates could be angles, distances between objects, etc.

Velocities will be represented through $\mathbf{v}_i \equiv \dot{\mathbf{r}}_i$ for spatial coordinates, or through \dot{q}_i for generalized coordinates. Note that dots above a symbol will always denote the total time derivative $\frac{d}{dt}$. Momenta are likewise either Newtonian momenta $\mathbf{p}_i = m_i \mathbf{v}_i$ or generalized momenta p_i .

For a fixed set of masses m_i Newton's 2nd law can be expressed in 2 equivalent ways:

1. It can be expressed as N second-order equations $\mathbf{F}_i = \frac{d}{dt}(m_i \dot{\mathbf{r}}_i)$ with $2N$ boundary conditions given in $\mathbf{r}_i(0)$ and $\dot{\mathbf{r}}_i(0)$. The problem then becomes one of determining the N vector variables $\mathbf{r}_i(t)$.
2. It can also be expressed as an equivalent set of $2N$ 1st order equations $\mathbf{F}_i = \dot{\mathbf{p}}_i$ & $\mathbf{p}_i/m_i = \dot{\mathbf{r}}_i$ with $2N$ boundary conditions given in $\mathbf{r}_i(0)$ and $\mathbf{p}_i(0)$. The problem then becomes one of determining the $2N$ vector variables $\mathbf{r}_i(t)$ and $\mathbf{p}_i(t)$.

Note that $\mathbf{F} = m\mathbf{a}$ holds in *inertial frames*. Inertial frames describe time and space homogeneously (invariant to displacements), isotropically (invariant to rotations), and in a time independent manner. These are frames where the motion of a particle not subject to any force is in a straight line with constant velocity. (The converse, that straight line motion with constant velocity implies there are no forces present, is not true. For example, there may be forces that are not visible for motion in a particular direction, like motion parallel to a constant magnetic field.) Noninertial frames also generically have “fictitious forces”, such as the centrifugal and Coriolis effects. Inertial frames also play a key role in special relativity. In general relativity the concept of inertial frames is replaced by that of geodesic motion.

The existence of an inertial frame is a useful approximation for working out the dynamics of particles, and non-inertial terms can often be included as perturbative corrections. Examples of approximate inertial frames are that of a fixed Earth, or better yet, of fixed distant stars. We can still test for how noninertial we are by looking for fictitious forces that (a) may point back to an origin with no source for the force or (b) behave in a non-standard fashion in different frames (i.e. they transform in an unusual manner when going between different frames).

We will use primes to denote coordinate transformations. If \mathbf{r} is measured in an inertial frame S , and \mathbf{r}' is measured in frame S' whose relation to S is given by a transformation $\mathbf{r}' = \mathbf{f}(\mathbf{r}, t)$, then S' is inertial iff $\ddot{\mathbf{r}} = 0 \leftrightarrow \ddot{\mathbf{r}}' = 0$. This is solved by the Galilean transformations,

$$\begin{aligned}\mathbf{r}' &= \mathbf{r} + \mathbf{v}_0 t, \\ t' &= t,\end{aligned}$$

with constant \mathbf{v}_0 . This transformation preserves the inertiality of frames, with $\mathbf{F} = m\ddot{\mathbf{r}}$ equivalent to $\mathbf{F}' = m\ddot{\mathbf{r}}'$. Galilean transformations are the non-relativistic limit, $v \ll c$, of Lorentz transformations which preserve inertial frames in special relativity. A few examples related to the concepts of inertial frames are:

1. In a rotating frame, the transformation is given by

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

If $\theta = \omega t$ for some constant ω , then $\ddot{\mathbf{r}} = 0$ does not imply $\{\ddot{x}', \ddot{y}'\} = 0$, so the primed frame is noninertial.

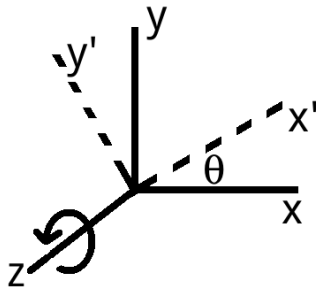


Figure 1.1: Frame rotated by an angle θ

2. In polar coordinates, $\mathbf{r} = r\hat{r}$, and the unit vectors are time dependent,

$$\frac{d\hat{r}}{dt} = \dot{\theta}\hat{\theta}, \quad \frac{d\hat{\theta}}{dt} = -\dot{\theta}\hat{r} \quad (1.1)$$

and thus

$$\ddot{\mathbf{r}} = \ddot{r}\hat{r} + 2\dot{r}\dot{\theta}\hat{\theta} + r(\ddot{\theta}\hat{\theta} - \dot{\theta}^2\hat{r}). \quad (1.2)$$

Even if $\ddot{\mathbf{r}} = 0$ we can still have $\ddot{r} \neq 0$ and $\ddot{\theta} \neq 0$, and we can not in general form a simple Newtonian force law equation $m\ddot{q} = F_q$ for each of these coordinates. This is different than the first example, since here we are picking coordinates rather than changing the reference frame, so to remind ourselves about their behavior we will call these “non-inertial coordinates” (which we may for example decide to use in an inertial frame). In general, curvilinear coordinates are non-inertial.

Lagrangian Mechanics

In Lagrangian mechanics, the key function is the Lagrangian

$$L = L(q, \dot{q}, t). \quad (1.3)$$

Here, $q = (q_1, \dots, q_N)$ and likewise $\dot{q} = (\dot{q}_1, \dots, \dot{q}_N)$. We are now letting N denote the number of scalar (rather than vector) variables, and will often use the short form to denote

dependence on these variables, as in Eq. (1.3). Typically we can write $L = T - V$ where T is the kinetic energy and V is the potential energy. In the simplest cases, $T = T(\dot{q})$ and $V = V(q)$, but we also allow the more general possibility that $T = T(q, \dot{q}, t)$ and $V = V(q, \dot{q}, t)$. It turns out, as we will discuss later, that even this generalization does not describe all possible classical mechanics problems.

The solution to a given mechanical problem is obtained by solving a set of N second-order differential equations known as *Euler-Lagrange equations of motion*,

$$\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_i} \right) - \frac{\partial L}{\partial q_i} = 0. \quad (1.4)$$

These equations involve \ddot{q}_i , and reproduce the Newtonian equations $\mathbf{F} = m\mathbf{a}$. The principle of stationary action (Hamilton's principle),

$$\delta S = \delta \int_{t_1}^{t_2} L(q, \dot{q}, t) dt = 0, \quad (1.5)$$

is the starting point for deriving the Euler-Lagrange equations. Although you have covered the Calculus of Variations in an earlier course on Classical Mechanics, we will review the main ideas in Section 1.5.

There are several advantages to working with the Lagrangian formulation, including

1. It is easier to work with the scalars T and V rather than vectors like \mathbf{F} .
2. The same formula in equation (1.4) holds true regardless of the choice of coordinates. To demonstrate this, let us consider new coordinates

$$Q_i = Q_i(q_1, \dots, q_N, t). \quad (1.6)$$

This particular sort of transformation is called a *point transformation*. Defining the new Lagrangian by

$$L' = L'(Q, \dot{Q}, t) = L(q, \dot{q}, t), \quad (1.7)$$

we claim that the equations of motion are simply

$$\frac{d}{dt} \left(\frac{\partial L'}{\partial \dot{Q}_i} \right) - \frac{\partial L'}{\partial Q_i} = 0. \quad (1.8)$$

Proof: (for $N = 1$, since the generalization is straightforward)

Given $L'(Q, \dot{Q}, t) = L(q, \dot{q}, t)$ with $Q = Q(q, t)$ then

$$\dot{Q} = \frac{d}{dt} Q(q, t) = \frac{\partial Q}{\partial q} \dot{q} + \frac{\partial Q}{\partial t}. \quad (1.9)$$

Therefore

$$\frac{\partial \dot{Q}}{\partial \dot{q}} = \frac{\partial Q}{\partial q}, \quad (1.10)$$

a result that we will use again in the future. Then

$$\begin{aligned} \frac{\partial L}{\partial q} &= \frac{\partial L'}{\partial q} = \frac{\partial L'}{\partial Q} \frac{\partial Q}{\partial q} + \frac{\partial L'}{\partial \dot{Q}} \frac{\partial \dot{Q}}{\partial q}, \\ \frac{\partial L}{\partial \dot{q}} &= \frac{\partial L'}{\partial \dot{q}} = \frac{\partial L'}{\partial \dot{Q}} \frac{\partial \dot{Q}}{\partial \dot{q}} = \frac{\partial L'}{\partial \dot{Q}} \frac{\partial Q}{\partial q}. \end{aligned} \quad (1.11)$$

Since $\frac{\partial Q}{\partial \dot{q}} = 0$ there is no term $\frac{\partial L'}{\partial Q} \frac{\partial Q}{\partial \dot{q}}$ in the last line.

Plugging these results into $0 = \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}} \right) - \frac{\partial L}{\partial q}$ gives

$$\begin{aligned} 0 &= \left[\frac{d}{dt} \left(\frac{\partial L'}{\partial \dot{Q}} \right) \frac{\partial Q}{\partial q} + \frac{\partial L'}{\partial \dot{Q}} \frac{d}{dt} \left(\frac{\partial Q}{\partial q} \right) \right] - \left[\frac{\partial L'}{\partial Q} \frac{\partial Q}{\partial q} + \frac{\partial L'}{\partial \dot{Q}} \frac{\partial \dot{Q}}{\partial q} \right] \\ &= \left[\frac{d}{dt} \left(\frac{\partial L'}{\partial \dot{Q}} \right) - \frac{\partial L'}{\partial Q} \right] \frac{\partial Q}{\partial q}. \end{aligned} \quad (1.12)$$

Here we have used $\frac{d}{dt} \frac{\partial Q}{\partial q} = (\dot{q} \frac{\partial}{\partial q} + \frac{\partial}{\partial t}) \frac{\partial Q}{\partial q} = \frac{\partial}{\partial q} (\dot{q} \frac{\partial}{\partial q} + \frac{\partial}{\partial t}) Q = \frac{\partial \dot{Q}}{\partial q}$ so that the second and fourth terms in the first line of Eq. (1.12) cancel. Finally for non-trivial transformation where $\frac{\partial Q}{\partial q} \neq 0$ we have, as expected,

$$0 = \frac{d}{dt} \left(\frac{\partial L'}{\partial \dot{Q}} \right) - \frac{\partial L'}{\partial Q}. \quad (1.13)$$

Note two things:

- This implies we can freely use the Euler-Lagrange equations for noninertial coordinates.
 - We can formulate L in whatever coordinates are easiest, and then change to convenient variables that better describe the symmetry of a system (for example, Cartesian to spherical).
3. Continuing our list of advantages for using L , we note that it is also easy to incorporate *constraints*. Examples include a mass constrained to a surface or a disk rolling without slipping. Often when using L we can avoid discussing forces of constraint (for example, the force normal to the surface).

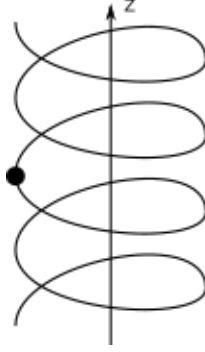


Figure 1.2: Bead on a helix

Lets discuss the last point in more detail (we will also continue to discuss it in the next section). The method for many problems with constraints is to simply make a good choice for the generalized coordinates to use for the Lagrangian, picking $N - k$ independent variables q_i for a system with k constraints.

Example: For a bead on a helix as in Fig. 1.2 we only need one variable, $q_1 = z$.

Example: A mass m_2 attached by a massless pendulum to a horizontally sliding mass m_1 as in Fig. 1.3, can be described with two variables $q_1 = x$ and $q_2 = \theta$.

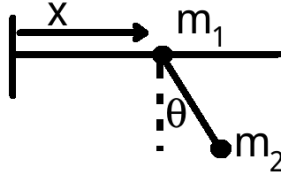


Figure 1.3: Pendulum of mass m_2 hanging on a rigid bar of length ℓ whose support m_1 is a frictionless horizontally sliding bead

Example: As an example using non-inertial coordinates consider a potential $V = V(r, \theta)$ in polar coordinates for a fixed mass m at position $\mathbf{r} = r\hat{r}$. Since $\dot{\mathbf{r}} = \dot{r}\hat{r} + r\dot{\theta}\hat{\theta}$ we have $T = \frac{m}{2}\dot{\mathbf{r}}^2 = \frac{m}{2}(\dot{r}^2 + r^2\dot{\theta}^2)$, giving

$$L = \frac{m}{2} (\dot{r}^2 + r^2\dot{\theta}^2) - V(r, \theta). \quad (1.14)$$

For r the Euler-Lagrange equation is

$$0 = \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{r}} \right) - \frac{\partial L}{\partial r} = \frac{d}{dt} (m\dot{r}) - mr\dot{\theta}^2 + \frac{\partial V}{\partial r}. \quad (1.15)$$

This gives

$$m\ddot{r} - mr\dot{\theta}^2 = -\frac{\partial V}{\partial r} = F_r, \quad (1.16)$$

from which we see that $F_r \neq m\ddot{r}$. For θ the Euler-Lagrange equation is

$$0 = \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{\theta}} \right) - \frac{\partial L}{\partial \theta} = \frac{d}{dt} (mr^2\dot{\theta}) + \frac{\partial V}{\partial \theta}. \quad (1.17)$$

This gives

$$\frac{d}{dt} (mr^2\dot{\theta}) = -\frac{\partial V}{\partial \theta} = F_\theta, \quad (1.18)$$

which is equivalent to the relation between angular momentum and torque perpendicular to the plane, $\dot{L}_z = F_\theta = \tau_z$. (Recall $\mathbf{L} = \mathbf{r} \times \mathbf{p}$ and $\boldsymbol{\tau} = \mathbf{r} \times \mathbf{F}$.)

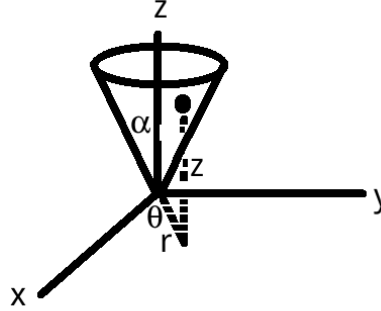


Figure 1.4: Particle on the inside of a cone

Example: Let us consider a particle rolling due to gravity in a frictionless cone, shown in Fig. 1.4, whose opening angle α defines an equation for points on the cone $\tan(\alpha) = \sqrt{x^2 + y^2}/z$. There are 4 steps which we can take to solve this problem (which are more general than this example):

1. Formulate T and V by $N = 3$ generalized coordinates. Here it is most convenient to choose cylindrical coordinates denoted (r, θ, z) , so that $T = \frac{m}{2} (\dot{r}^2 + r^2\dot{\theta}^2 + \dot{z}^2)$ and $V = mgz$.
2. Reduce the problem to $N - k = 2$ independent coordinates and determine the new Lagrangian $L = T - V$. In this case we eliminate $z = r \cot(\alpha)$ and $\dot{z} = \dot{r} \cot(\alpha)$, so

$$L = \frac{m}{2} \left[(1 + \cot^2 \alpha) \dot{r}^2 + r^2 \dot{\theta}^2 \right] - mgr \cot \alpha. \quad (1.19)$$

3. Find the Euler-Lagrange equations. For r , $0 = \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{r}} \right) - \frac{\partial L}{\partial r}$, which here is

$$0 = \frac{d}{dt} [m(1 + \cot^2 \alpha) \dot{r}] - mr\dot{\theta}^2 + mg \cot \alpha \quad (1.20)$$

giving

$$(1 + \cot^2 \alpha) \ddot{r} - r\dot{\theta}^2 + g \cot \alpha = 0. \quad (1.21)$$

For θ we have $0 = \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{\theta}} \right) - \frac{\partial L}{\partial \theta}$, so

$$0 = \frac{d}{dt} (mr^2\dot{\theta}) - 0, \quad (1.22)$$

giving

$$(2\dot{r}\dot{\theta} + r\ddot{\theta})r = 0. \quad (1.23)$$

4. Solve the system analytically or numerically, for example using Mathematica. Or we might be only interested in determining certain properties or characteristics of the motion without a full solution (we will discuss this last point in detail in Ch.7).

Hamiltonian Mechanics

In Hamiltonian mechanics, the canonical momenta $p_i \equiv \frac{\partial L}{\partial \dot{q}_i}$ are promoted to coordinates on equal footing with the generalized coordinates q_i . The coordinates (q, p) are *canonical variables*, and the space of canonical variables is known as *phase space*.

The Euler-Lagrange equations say $\dot{p}_i = \frac{\partial L}{\partial q_i}$. These need not equal the kinematic momenta $m_i \dot{q}_i$ if $V = V(q, \dot{q})$. Performing the Legendre transformation

$$H(q, p, t) = \dot{q}_i p_i - L(q, \dot{q}, t) \quad (1.24)$$

(where for this equation, and henceforth, repeated indices will imply a sum unless otherwise specified) yields the *Hamilton equations of motion*

$$\begin{aligned} \dot{q}_i &= \frac{\partial H}{\partial p_i} \\ \dot{p}_i &= -\frac{\partial H}{\partial q_i} \end{aligned} \quad (1.25)$$

which are $2N$ 1st order equations. We also have the result that

$$\frac{\partial H}{\partial t} = -\frac{\partial L}{\partial t}. \quad (1.26)$$

Proof: (for $N = 1$) Consider how the left and right hand side of the Legendre transformation behave under differential changes to the variables,

$$\begin{aligned} dH &= \frac{\partial H}{\partial q} dq + \frac{\partial H}{\partial p} dp + \frac{\partial H}{\partial t} dt \\ &= \cancel{p d\dot{q}} + \dot{q} dp - \frac{\partial L}{\partial q} dq - \cancel{\frac{\partial L}{\partial \dot{q}} d\dot{q}} - \frac{\partial L}{\partial t} dt = \dot{q} dp - \dot{p} dq - \frac{\partial L}{\partial t} dt. \end{aligned} \quad (1.27)$$

In the last line we used the definition $p = \frac{\partial L}{\partial \dot{q}}$ and the E-L equation, $\dot{p} = \frac{\partial L}{\partial q}$. Since we are free to independently vary dq , dp , and dt this implies (generalizing now to N variables):

$$-\frac{\partial H}{\partial q_i} = \dot{p}_i, \quad \frac{\partial H}{\partial p_i} = \dot{q}_i, \quad \frac{\partial H}{\partial t} = -\frac{\partial L}{\partial t}. \quad (1.28)$$

The first two results are Hamilton's equations. We can interpret these two Hamilton equations as follows:

- $\dot{q}_i = \frac{\partial H(q,p,t)}{\partial p_i} = \dot{q}_i(q,p,t)$ is an inversion of $p_i = \frac{\partial L(q,\dot{q},t)}{\partial \dot{q}_i} = p_i(q,\dot{q},t)$.
- $\dot{p}_i = -\frac{\partial H}{\partial q_i}$ provides the Newtonian dynamics.

However, these two equations have an equal footing in Hamiltonian mechanics, since the coordinates and momenta are treated on a common ground. We can use $p_i = \frac{\partial L}{\partial \dot{q}_i}$ to construct H from L and then forget about L .

As an example of the manner in which we will usually consider transformations between Lagrangians and Hamiltonians, consider again the variables relevant for the particle on a cone from Fig. 1.4:

$$\begin{array}{ccccc} L(r, \theta, z, \dot{r}, \dot{\theta}, \dot{z}) & \xrightarrow{z=r \cot \alpha} & \text{new } L(r, \theta, \dot{r}, \dot{\theta}) & \longrightarrow & \text{Euler-Lagrange Eqns.} \\ \updownarrow & & \updownarrow & & \updownarrow \\ H(r, \theta, z, p_r, p_\theta, p_z) & \xRightarrow{\text{not here}} & H(r, \theta, p_r, p_\theta) & \longrightarrow & \text{Hamilton Eqns.} \end{array} \quad (1.29)$$

Here we consider transforming between L and H either before or after removing the redundant coordinate z , but in this course we will only consider constraints imposed on Lagrangians and not in the Hamiltonian formalism (the step indicated by \Rightarrow). For the curious, the topic of imposing constraints on Hamiltonians, including even more general constraints than those we will consider, is covered well in Dirac's little book "Lectures on Quantum Mechanics". Although Hamiltonian and Lagrangian mechanics provide equivalent formalisms, there is often an advantage to using one or the other. In the case of Hamiltonian mechanics potential advantages include the language of phase space with Liouville's Theorem, Poisson Brackets and the connection to quantum mechanics, as well as the Hamilton-Jacobi transformation theory (all to be covered later on).

Special case: Let us consider a special case that is sufficient to imply that the Hamiltonian is equal to the energy, $H = E \equiv T + V$. If we only have quadratic dependence on velocities in the kinetic energy, $T = \frac{1}{2} T_{jk}(q) \dot{q}_j \dot{q}_k$, and $V = V(q)$ with $L = T - V$, then

$$\dot{q}_i p_i = \dot{q}_i \frac{\partial L}{\partial \dot{q}_i} = \frac{1}{2} \dot{q}_i T_{ik} \dot{q}_k + \frac{1}{2} \dot{q}_j T_{ji} \dot{q}_i = 2T. \quad (1.30)$$

Hence,

$$H = \dot{q}_i p_i - L = T + V = E \quad (1.31)$$

which is just the energy.

Another Special case: Consider a class of Lagrangians given as

$$L(q, \dot{q}, t) = L_0 + a_j \dot{q}_j + \frac{1}{2} \dot{q}_j T_{jk} \dot{q}_k \quad (1.32)$$

where $L_0 = L_0(q, t)$, $a_j = a_j(q, t)$, and $T_{jk} = T_{kj} = T_{jk}(q, t)$. We can write this in shorthand as

$$L = L_0 + \vec{a} \cdot \dot{\vec{q}} + \frac{1}{2} \dot{\vec{q}} \cdot \hat{T} \cdot \dot{\vec{q}}. \quad (1.33)$$

Here the generalized coordinates, momenta, and coefficients have been collapsed into vectors, like \vec{q} (rather than the boldface that we reserve for Cartesian vectors), and dot products of vectors from the left imply transposition of that vector. Note that \vec{q} is an unusual vector, since its components can have different dimensions, eg. $\vec{q} = (x, \theta)$, but nevertheless this notation is useful. To find H ,

$$p_j = \frac{\partial L}{\partial \dot{q}_j} = a_j + T_{jk} \dot{q}_k, \quad (1.34)$$

meaning $\vec{p} = \vec{a} + \hat{T} \cdot \dot{\vec{q}}$. Inverting this gives $\dot{\vec{q}} = \hat{T}^{-1} \cdot (\vec{p} - \vec{a})$, where \hat{T}^{-1} will exist because of the positive-definite nature of kinetic energy, which implies that \hat{T} is a positive definite matrix. Thus, $H = \dot{\vec{q}} \cdot \vec{p} - L$ yields

$$H = \frac{1}{2} (\vec{p} - \vec{a}) \cdot \hat{T}^{-1} \cdot (\vec{p} - \vec{a}) - L_0(q, t) \quad (1.35)$$

as the Hamiltonian. So for *any* Lagrangian in the form of Eq. (1.32), we can find \hat{T}^{-1} and write down the Hamiltonian as in Eq. (1.35) immediately.

Example: let us consider $L = \frac{1}{2} m \mathbf{v}^2 - e\phi + e\mathbf{A} \cdot \mathbf{v}$, where e is the electric charge and SI units are used. In Eq. (1.32), because the coordinates are Cartesian, $\mathbf{a} = e\mathbf{A}$, $\hat{T} = m\mathbb{1}$, and $L_0 = -e\phi$, so

$$H = \frac{1}{2m} (\mathbf{p} - e\mathbf{A})^2 + e\phi. \quad (1.36)$$

As you have presumably seen in an earlier course, this Hamiltonian does indeed reproduce the Lorentz force equation $e(\mathbf{E} + \mathbf{v} \times \mathbf{B}) = m\dot{\mathbf{v}}$.

A more detailed Example. Find L and H for the frictionless pendulum shown in Fig. 1.3. This system has two constraints, that m_1 is restricted to lie on the x-axis sliding without friction, and that the rod between m_1 and m_2 is rigid, giving

$$y_1 = 0, \quad (y_1 - y_2)^2 + (x_1 - x_2)^2 = \ell^2. \quad (1.37)$$

Prior to imposing any constraints the Lagrangian is

$$L = T - V = \frac{m_1}{2}\dot{x}_1^2 + \frac{m_2}{2}(\dot{x}_2^2 + \dot{y}_2^2) - m_2gy_2 - m_1gy_1. \quad (1.38)$$

Lets choose to use $x \equiv x_1$ and the angle θ as the independent coordinates after imposing the constraints in Eq. (1.37). This allows us to eliminate $y_1 = 0$, $x_2 = x + \ell \sin \theta$ and $y_2 = -\ell \cos \theta$, together with $\dot{x}_2 = \dot{x} + \ell \cos \theta \dot{\theta}$, $\dot{y}_2 = \ell \sin \theta \dot{\theta}$, $\dot{x}_1 = \dot{x}$. The Lagrangian with constraints imposed is

$$L = \frac{m_1}{2}\dot{x}^2 + \frac{m_2}{2}(\dot{x}^2 + 2\ell \cos \theta \dot{x} \dot{\theta} + \ell^2 \cos^2 \theta \dot{\theta}^2 + \ell^2 \sin^2 \theta \dot{\theta}^2) + m_2g\ell \cos \theta. \quad (1.39)$$

Next we determine the Hamiltonian. First we find

$$\begin{aligned} p_x &= \frac{\partial L}{\partial \dot{x}} = m_1\dot{x} + m_2(\dot{x} + \ell \cos \theta \dot{\theta}) = (m_1 + m_2)\dot{x} + m_2\ell \cos \theta \dot{\theta}, \\ p_\theta &= \frac{\partial L}{\partial \dot{\theta}} = m_2\ell \cos \theta \dot{x} + m_2\ell^2 \dot{\theta}. \end{aligned} \quad (1.40)$$

Note that p_x is not simply proportional to \dot{x} here (actually p_x is the center-of-mass momentum). Writing $\begin{pmatrix} p_x \\ p_\theta \end{pmatrix} = \hat{T} \cdot \begin{pmatrix} \dot{x} \\ \dot{\theta} \end{pmatrix}$ gives

$$\hat{T} = \begin{pmatrix} m_1 + m_2 & m_2\ell \cos \theta \\ m_2\ell \cos \theta & m_2\ell^2 \end{pmatrix}, \quad (1.41)$$

with $L = \frac{1}{2}(\dot{x} \ \dot{\theta}) \cdot \hat{T} \cdot \begin{pmatrix} \dot{x} \\ \dot{\theta} \end{pmatrix} + L_0$ where $L_0 = m_2g\ell \cos \theta$. Computing

$$\hat{T}^{-1} = \frac{1}{m_1m_2\ell^2 + m_2\ell^2 \sin^2 \theta} \begin{pmatrix} m_2\ell^2 & -m_2\ell \cos \theta \\ -m_2\ell \cos \theta & m_1 + m_2 \end{pmatrix}, \quad (1.42)$$

we can simply apply Eq. (1.35) to find the corresponding Hamiltonian

$$\begin{aligned} H &= \frac{1}{2}(p_x \ p_\theta) \cdot \hat{T}^{-1} \cdot \begin{pmatrix} p_x \\ p_\theta \end{pmatrix} - m_2g\ell \cos \theta \\ &= \frac{1}{2m_2\ell^2(m_1 + m_2 \sin^2 \theta)} \left[m_2\ell^2 p_x^2 + (m_1 + m_2)p_\theta^2 - 2m_2\ell \cos \theta p_x p_\theta \right] - m_2g\ell \cos \theta. \end{aligned} \quad (1.43)$$

Lets compute the Hamilton equations of motion for this system. First for (x, p_x) we find

$$\begin{aligned} \dot{x} &= \frac{\partial H}{\partial p_x} = \frac{p_x}{m_1 + m_2 \sin^2 \theta} - \frac{\cos \theta p_\theta}{\ell(m_1 + m_2 \sin^2 \theta)}, \\ \dot{p}_x &= -\frac{\partial H}{\partial x} = 0. \end{aligned} \quad (1.44)$$

As we might expect, the CM momentum is time independent. Next for (θ, p_θ) :

$$\begin{aligned}\dot{\theta} &= \frac{\partial H}{\partial p_\theta} = \frac{1}{m_2 \ell^2 (m_1 + m_2 \sin^2 \theta)} \left[(m_1 + m_2) p_\theta - m_2 \ell \cos \theta p_x \right], \\ \dot{p}_\theta &= -\frac{\partial H}{\partial \theta} = \frac{\sin \theta \cos \theta}{\ell^2 (m_1 + m_2 \sin^2 \theta)^2} \left[m_2 \ell^2 p_x^2 + (m_1 + m_2) p_\theta^2 - 2 m_2 \ell \cos \theta p_x p_\theta \right] \\ &\quad - m_2 g \ell \sin \theta - \frac{\sin \theta p_x p_\theta}{\ell (m_1 + m_2 \sin \theta)}.\end{aligned}\tag{1.45}$$

These non-linear coupled equations are quite complicated, but could be solved in mathematica or another numerical package. To test our results for these equations of motion analytically, we can take the small angle limit, approximating $\sin \theta \simeq \theta$, $\cos \theta \simeq 1$ to obtain

$$\begin{aligned}\dot{x} &= \frac{p_x}{m_1} - \frac{p_\theta}{\ell m_1}, \quad \dot{p}_x = 0, \quad \dot{\theta} = \frac{1}{m_1 m_2 \ell^2} \left[(m_1 + m_2) p_\theta - m_2 \ell p_x \right], \\ \dot{p}_\theta &= \frac{\theta}{\ell^2 m_1^2} \left[m_2 \ell^2 p_x^2 + (m_1 + m_2) p_\theta^2 - 2 m_2 \ell \cos \theta p_x p_\theta \right] - \frac{\theta p_x p_\theta}{\ell m_1} - m_2 g \ell \theta.\end{aligned}\tag{1.46}$$

To simplify it further we can work in the CM frame, thus setting $p_x = 0$, and linearize the equations by noting that $p_\theta \sim \dot{\theta}$ should be small for θ to remain small, and hence θp_θ^2 is a higher order term. For the non-trivial equations this leaves

$$\dot{x} = -\frac{p_\theta}{\ell m_1}, \quad \dot{\theta} = \frac{p_\theta}{\mu \ell^2}, \quad \dot{p}_\theta = -m_2 g \ell \theta,\tag{1.47}$$

where $\mu = m_1 m_2 / (m_1 + m_2)$ is the reduced mass for the two-body system. Thus $\ddot{\theta} = \dot{p}_\theta / (\mu \ell^2) = -\frac{m_2 g}{\mu \ell} \theta$ as expected for simple harmonic motion.

1.3 Symmetry and Conservation Laws

A *cyclic coordinate* is one which does not appear in the Lagrangian, or equivalently in the Hamiltonian. Because $H(q, p, t) = \dot{q}_i p_i - L(q, \dot{q}, t)$, if q_j is absent in L for some particular j , it will be absent in H as well. The absence of that q_j corresponds with a symmetry in the dynamics.

In this context, *Noether's theorem* means that a symmetry implies a cyclic coordinate, which in turn produces a conservation law. If q_j is a cyclic coordinate for some j , then we can change that coordinate without changing the dynamics given by the Lagrangian or Hamiltonian, and hence there is a symmetry. Furthermore the corresponding canonical momentum p_j is conserved, meaning it is a constant through time.

The proof is simple. If $\frac{\partial L}{\partial q_j} = 0$ then $\dot{p}_j = \frac{d}{dt} \frac{\partial L}{\partial \dot{q}_j} = \frac{\partial L}{\partial q_j} = 0$, or even more simply, $\frac{\partial H}{\partial q_j} = 0$ is equivalent to $\dot{p}_j = 0$, so p_j is a constant in time.

Special cases and examples of this abound. Lets consider a few important ones:

1. Consider a system of N particles where no external or internal force acts on the center of mass (CM) coordinate $\mathbf{R} = \frac{1}{M} \sum_i m_i \mathbf{r}_i$, where the total mass $M = \sum_i m_i$. Then the CM momentum \mathbf{P} is conserved. This is because

$$\mathbf{F}_{\mathbf{R}} = -\nabla_{\mathbf{R}} V = 0 \quad (1.48)$$

so V is independent of \mathbf{R} . Meanwhile, $T = \frac{1}{2} \sum_i m_i \dot{\mathbf{r}}_i^2$, which when using coordinates relative to the center of mass, $\mathbf{r}'_i \equiv \mathbf{r}_i - \mathbf{R}$, becomes

$$T = \frac{1}{2} \left(\sum_i m_i \right) \dot{\mathbf{R}}^2 + \dot{\mathbf{R}} \cdot \frac{d}{dt} \left(\sum_i m_i \mathbf{r}'_i \right) + \frac{1}{2} \sum_i m_i \dot{\mathbf{r}}_i'^2 = \frac{1}{2} M \dot{\mathbf{R}}^2 + \frac{1}{2} \sum_i m_i \dot{\mathbf{r}}_i'^2. \quad (1.49)$$

Note that $\sum_i m_i \mathbf{r}'_i = 0$ from the definitions of M , \mathbf{R} , and \mathbf{r}'_i , so T splits into two terms, one for the CM motion and one for relative motion. We also observe that T is independent of \mathbf{R} . This means that \mathbf{R} is cyclic for the full Lagrangian L , so $\mathbf{P} = M \dot{\mathbf{R}}$ is a conserved quantity. In our study of rigid bodies we will also need the forms of M and \mathbf{R} for a continuous body with mass distribution $\rho(\mathbf{r})$, which for a three dimensional body are $M = \int d^3r \rho(\mathbf{r})$ and $\mathbf{R} = \frac{1}{M} \int d^3r \rho(\mathbf{r}) \mathbf{r}$.

Note that $\dot{\mathbf{P}} = 0$ is satisfied by having no total external force, so $\mathbf{F}^{\text{ext}} = \sum_i \mathbf{F}_i^{\text{ext}} = 0$, and by the internal forces obeying Newton's 3rd law $\mathbf{F}_{i \rightarrow j} = -\mathbf{F}_{j \rightarrow i}$. Hence,

$$M \ddot{\mathbf{R}} = \sum_i \mathbf{F}_i^{\text{ext}} + \sum_{i,j} \mathbf{F}_{i \rightarrow j} = 0. \quad (1.50)$$

2. Let us consider a system that is invariant with respect to rotations of angle ϕ about a symmetry axis. This has a conserved angular momentum. If we pick ϕ as a generalized coordinate, then $L = T - V$ is independent of ϕ , so $\dot{p}_\phi = \frac{\partial L}{\partial \phi} = 0$ meaning p_ϕ is constant. In particular, for a system where V is independent of the angular velocity $\dot{\phi}$ we have

$$p_\phi = \frac{\partial T}{\partial \dot{\phi}} = \sum_i m_i \dot{\mathbf{r}}_i \cdot \frac{\partial \dot{\mathbf{r}}_i}{\partial \dot{\phi}} = \sum_i m_i \mathbf{v}_i \cdot \frac{\partial \mathbf{r}_i}{\partial \phi}. \quad (1.51)$$

Simplifying further using the results in Fig. 2.2 yields

$$p_\phi = \sum_i m_i \mathbf{v}_i \cdot (\hat{n} \times \mathbf{r}_i) = \hat{n} \cdot \sum_i \mathbf{r}_i \times m_i \mathbf{v}_i = \hat{n} \cdot \mathbf{L}_{\text{total}}. \quad (1.52)$$

Note that \mathbf{L} about the CM is conserved for systems with no external torque, $\boldsymbol{\tau}^{\text{ext}} = \sum_i \mathbf{r}_i \times \mathbf{F}_i^{\text{ext}} = 0$ and internal forces that are all central. Defining $\mathbf{r}_{ij} \equiv \mathbf{r}_i - \mathbf{r}_j$ and its magnitude appropriately, this means $V_{ij} = V_{ij}(r_{ij})$. This implies that $\mathbf{F}_{ji} = -\nabla_i V_{ij}$ (no sum on the repeated index) is parallel to \mathbf{r}_{ij} . Hence,

$$\frac{d\mathbf{L}}{dt} = \sum_i \mathbf{r}_i \times \dot{\mathbf{p}}_i = \sum_i \mathbf{r}_i \times \mathbf{F}_i^{\text{ext}} + \sum_{i,j} \mathbf{r}_i \times \mathbf{F}_{ji}. \quad (1.53)$$

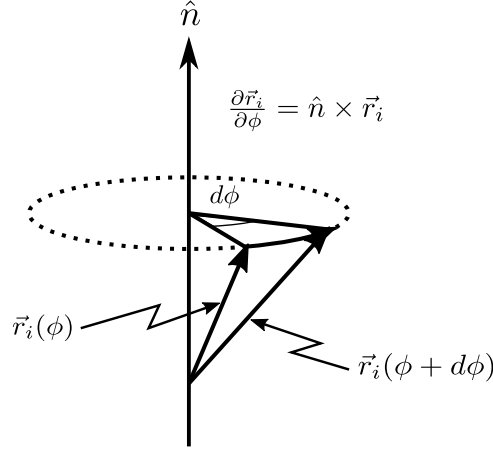


Figure 1.5: Rotation about a symmetry axis

However, $\sum_i \mathbf{r}_i \times \mathbf{F}_i^{\text{ext}} = 0$, so

$$\frac{d\mathbf{L}}{dt} = \sum_{i < j} \mathbf{r}_{ij} \times \mathbf{F}_{ji} = 0. \quad (1.54)$$

3. One can also consider a scaling transformation. Suppose that under the transformation $\mathbf{r}_i \rightarrow \lambda \mathbf{r}_i$ the potential is homogeneous and transforms as $V \rightarrow \lambda^k V$ for some constant k . Letting T be quadratic in $\dot{\mathbf{r}}_i$ and taking time to transform as $t \rightarrow \lambda^{1-k/2} t$ then gives $\dot{\mathbf{r}}_i \rightarrow \lambda^{k/2} \dot{\mathbf{r}}_i$. So by construction $T \rightarrow \lambda^k T$ also, and thus the full Lagrangian $L \rightarrow \lambda^k L$. This overall factor does not change the Euler-Lagrange equations, and hence the transformation is a symmetry of the dynamics, only changing the overall scale or units of the coordinate and time variables, but not their dynamical relationship. This can be applied for several well known potentials:
 - a) $k = 2$ for a harmonic oscillator. In this case the scaling for time is given by $1 - k/2 = 0$, so it does not change with λ . The amplitude for the coordinates scales with λ . Thus, the frequency of the oscillator, which is a time variable, is independent of the amplitude.
 - b) $k = -1$ for the Coulomb potential. Here $1 - k/2 = 3/2$ so there is a more intricate relation between coordinates and time. This power is consistent with the behavior of bound state orbits, where the period of the orbit T obeys $T^2 \propto a^3$, for a the semi-major axis distance (Kepler's 3rd law).
 - c) $k = 1$ for a uniform gravitational field. Here $1 - k/2 = 1/2$ so for a freely falling object, the time of free fall goes as \sqrt{h} where h is the distance fallen.
4. Consider the Lagrangian for a charge in the presence of scalar (ϕ) and vector (\mathbf{A}) electromagnetic fields, $L = \frac{1}{2} m \dot{\mathbf{r}}^2 - e\phi + e\mathbf{A} \cdot \dot{\mathbf{r}}$. As a concrete example, let us take ϕ and \mathbf{A} to be independent of the Cartesian coordinate x . The canonical momentum is

$\mathbf{p} = \frac{\partial L}{\partial \dot{\mathbf{r}}} = m\dot{\mathbf{r}} + e\mathbf{A}$, which is notably different from the kinetic momentum. Then x being cyclic means the canonical momentum p_x is conserved.

5. Let us consider the conservation of energy and the relationship between energy and the Hamiltonian. Applying the time derivative gives $\dot{H} = \frac{\partial H}{\partial q}\dot{q} + \frac{\partial H}{\partial p}\dot{p} + \frac{\partial H}{\partial t}$. However, $\dot{q} = \frac{\partial H}{\partial p}$ and $\dot{p} = -\frac{\partial H}{\partial q}$. Thus

$$\dot{H} = \frac{\partial H}{\partial t} = -\frac{\partial L}{\partial t}. \quad (1.55)$$

There are two things to consider.

- If H (or L) has no explicit time dependence, then $H = \dot{q}_i p_i - L$ is conserved.
- Energy is conserved if $\dot{E} = 0$, where energy is defined by $E = T + V$.

If $H = E$ then the two points are equivalent, but otherwise either of the two could be true while the other is false.

Example: Let us consider a system which provides an example where $H = E$ but energy is not conserved, and where $H \neq E$ but H is conserved. The two situations will be obtained from the same example by exploiting a coordinate choice. Consider a system consisting of a mass m attached by a spring of constant k to a cart moving at a constant speed v_0 in one dimension, as shown in Fig. 1.6. Let us call x the displacement

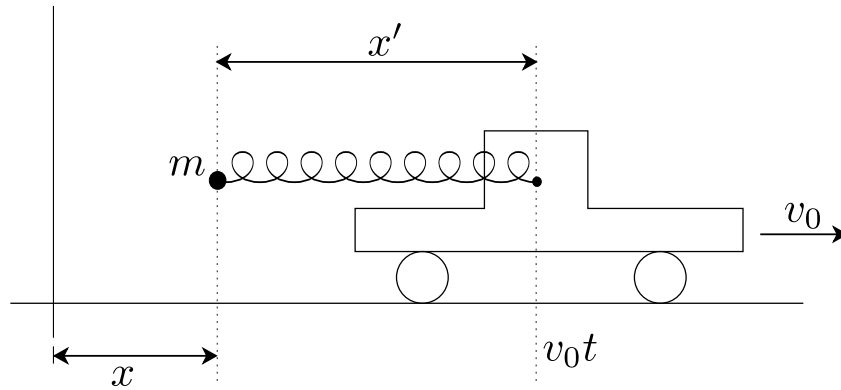


Figure 1.6: Mass attached by a spring to a moving cart

of m from the fixed wall and x' is its displacement from the center of the moving cart. Using x ,

$$L(x, \dot{x}) = T - V = \frac{m}{2}\dot{x}^2 - \frac{k}{2}(x - v_0 t)^2, \quad (1.56)$$

where the kinetic term is quadratic in \dot{x} and the potential term is independent of \dot{x} . This means that H falls in the special case considered in Eq. (1.31) so

$$H = E = T + V = \frac{p^2}{2m} + \frac{k}{2}(x - v_0 t)^2, \quad (1.57)$$

However $\frac{\partial H}{\partial t} \neq 0$ so the energy is not conserved. (Of course the full energy would be conserved, but we have not accounted for the energy needed to pull the cart at a constant velocity, treating that instead as external to our system. That is what led to the time dependent H .)

If we instead choose to use the coordinate $x' = x - v_0 t$, then

$$L'(x', \dot{x}') = \frac{m}{2}\dot{x}'^2 + mv_0 x' + \frac{m}{2}v_0^2 - \frac{k}{2}x'^2. \quad (1.58)$$

Note that $p' = m\dot{x}' + mv_0 = m\dot{x} = p$. This Lagrangian fits the general form in equation (1.32) with $a = mv_0$ and $L_0 = mv_0^2/2 - kx'^2/2$. So

$$H'(x', p') = \dot{x}'p' - L' = \frac{1}{2m}(p' - mv_0)^2 + \frac{k}{2}x'^2 - \frac{m}{2}v_0^2, \quad (1.59)$$

Here the last terms is a constant shift. The first and second terms in this expression for H' look kind of like the energy that we would calculate if we were sitting on the cart and did not know it was moving, which is not the same as the energy above. Hence, $H' \neq E$, but $\dot{H}' = 0$ because $\frac{\partial H'}{\partial t} = 0$, so H' is conserved.

1.4 Constraints and Friction Forces

So far, we've considered constraints to a surface or curve that are relationships between coordinates. These fall in the category of *holonomic constraints*. Such constraints take the form

$$f(q_1, \dots, q_N, t) = 0 \quad (1.60)$$

where explicit time dependence is allowed as a possibility. An example of holonomic constrain is mass in a cone (Figure 1.4), where the constrain is $z - r \cot \alpha = 0$. Constraints that violate the form in Eq. (1.60) are *non-holonomic constraints*.

- An example of a non-holonomic constraint is a mass on the surface of a sphere. The constraint here is an *inequality* $r^2 - a^2 \geq 0$ where r is the radial coordinate and a is the radius of the sphere.
- Another example of a non-holonomic constraint is an object rolling on a surface without slipping. The point of contact is stationary, so the constraint is actually on the *velocities*.

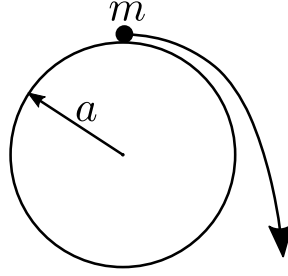


Figure 1.7: Mass on a sphere

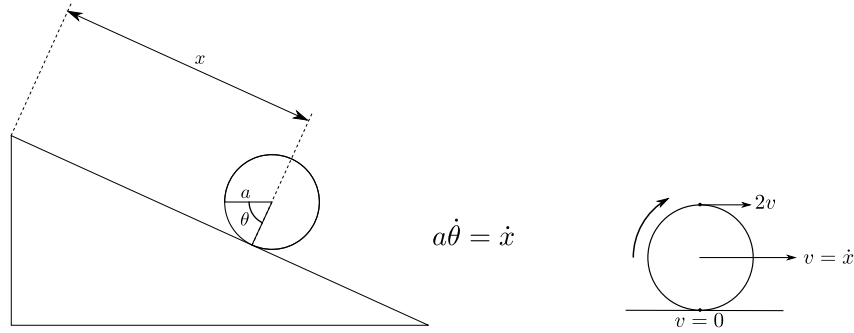


Figure 1.8: Disk rolling down an incline without slipping

A simple example is a disk of radius a rolling down an inclined plane without slipping, as shown in Fig. 1.8. Here the condition on velocities, $a\dot{\theta} = \dot{x}$ is simple enough that it can be integrated into a holonomic constraint.

As a more sophisticated example, consider a vertical disk of radius a rolling on a horizontal plane, as shown in Fig. 1.9. The coordinates are (x, y, θ, ϕ) , where (x, y) is the point of contact, ϕ is the rotation angle about its own axis, and θ is the angle of orientation along the xy -plane. We will assume that the flat edge of the disk always remains parallel to z , so the disk never tips over. The no-slip condition is $v = a\dot{\phi}$ where \mathbf{v} is the velocity of the center of the disk, and $v = |\mathbf{v}|$. This means $\dot{x} = v \sin(\theta) = a \sin(\theta)\dot{\phi}$ and $\dot{y} = -v \cos(\theta) = -a \cos(\theta)\dot{\phi}$, or in differential notation, $dx - a \sin(\theta)d\phi = 0$ and $dy + a \cos(\theta)d\phi = 0$.

In general, constraints of the form

$$\sum_j a_j(q) dq_j + a_t(q) dt = 0 \quad (1.61)$$

are not holonomic. We will call this a *semi-holonomic constraint*.

Let us consider the special case of a holonomic constraint in differential form, $f(q_1, \dots, q_{3N}, t) = 0$. This means

$$df = \sum_j \frac{\partial f}{\partial q_j} dq_j + \frac{\partial f}{\partial t} dt = 0, \quad (1.62)$$

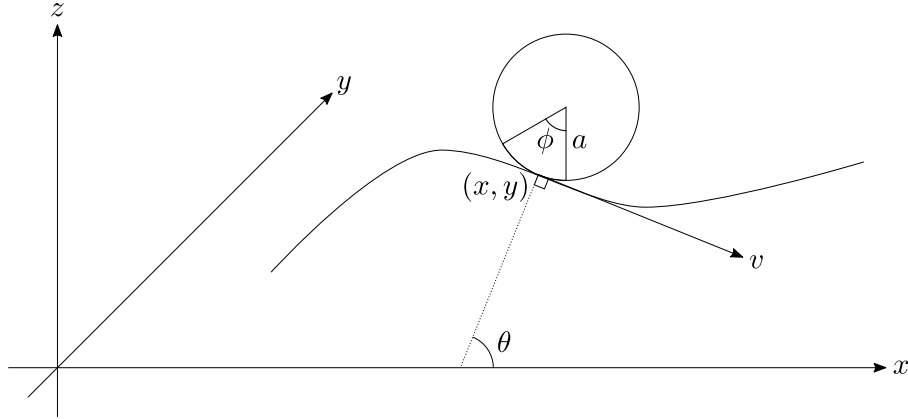


Figure 1.9: Vertical rolling disk on a two dimensional plane

so $a_j = \frac{\partial f}{\partial q_j}$ and $a_t = \frac{\partial f}{\partial t}$. The properties of mixed partial derivatives means

$$\frac{\partial a_j}{\partial q_i} = \frac{\partial a_i}{\partial q_j}, \quad \frac{\partial a_t}{\partial q_i} = \frac{\partial a_i}{\partial t}. \quad (1.63)$$

These conditions imply that a seemingly semi-holonomic constraint is in fact holonomic. (In math we would say that we have an exact differential form df for the holonomic case, but the differential form in Eq.(1.61) need not always be exact.)

Example: To demonstrate that not all semiholonomic constraints are secretly holonomic, consider the constraint in the example of the vertical disk. Here there is no function $h(x, y, \theta, \phi)$ that we can multiply the constraint $df = 0$ by to make it holonomic. For the vertical disk from before, we could try $(dx - a \sin(\theta) d\phi)h = 0$ with $a_x = h$, $a_\phi = -a \sin(\theta)h$, $a_\theta = 0$, and $a_y = 0$, all for some function h . As we must have $\frac{\partial a_\phi}{\partial \theta} = \frac{\partial a_\theta}{\partial \phi}$, then $0 = -a \cos(\theta) - a \sin(\theta) \frac{\partial h}{\partial \theta}$, so $h = \frac{k}{\sin(\theta)}$. That said, $\frac{\partial a_x}{\partial \theta} = \frac{\partial a_\theta}{\partial x}$ gives $\frac{\partial h}{\partial \theta} = 0$, yielding a contradiction for any non-trivial h with $k \neq 0$.

If the rolling is instead constrained to a line rather than a plane, then the constraint is holonomic. Take as an example $\theta = \frac{\pi}{2}$ for rolling along \hat{x} , then $\dot{x} = a\dot{\phi}$ and $\dot{y} = 0$. Integrating we have $x = a\phi + x_0$, $y = y_0$, and $\theta = \frac{\pi}{2}$, which together form a set of holonomic constraints.

A useful concept for discussing constraints is that of the *virtual displacement* $\delta \mathbf{r}_i$ of particle i . There are a few properties to be noted of $\delta \mathbf{r}_i$.

- It is infinitesimal.
- It is consistent with the constraints.
- It is carried out at a fixed time (so time dependent constraints do not change its form).

Example: let us consider a bead constrained to a moving wire. The wire is oriented along

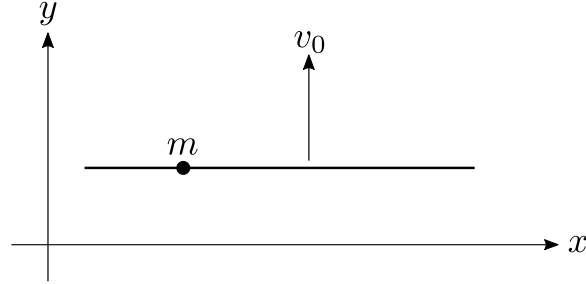


Figure 1.10: Bead on a moving wire

the x -axis and is moving with coordinate $y = v_0 t$. Here the virtual displacement of the bead $\delta \mathbf{r}$ is always parallel to \hat{x} (since it is determined at a fixed time), whereas the real displacement $d\mathbf{r}$ has a component along \hat{y} in a time interval dt .

For a large number of constraints, the constraint force \mathbf{Z}_i is perpendicular to $\delta \mathbf{r}_i$, meaning $\mathbf{Z}_i \cdot \delta \mathbf{r}_i = 0$, so the “virtual work” (in analogy to work $W = \int \mathbf{F} \cdot d\mathbf{r}$) of a constraint force vanishes. More generally, there is no *net* work from constraints, so $\sum_i \mathbf{Z}_i \cdot \delta \mathbf{r}_i = 0$ (which holds for the actions of surfaces, rolling constraints, and similar things). The Newtonian equation of motion is $\dot{\mathbf{p}}_i = \mathbf{F}_i + \mathbf{Z}_i$, where \mathbf{F}_i encapsulates other forces. Vanishing virtual work gives

$$\sum_i (\dot{\mathbf{p}}_i - \mathbf{F}_i) \cdot \delta \mathbf{r}_i = 0 \quad (1.64)$$

which is the *D’Alembert principle*. This could be taken as the starting principal for classical mechanics instead of the Hamilton principle of stationary action.

Of course Eq.(1.64) is not fully satisfactory since we are now used to the idea of working with generalized coordinates rather than the cartesian vector coordinates used there. So lets transform to generalized coordinates through $\mathbf{r}_i = \mathbf{r}_i(q, t)$, so $\delta \mathbf{r}_i = \frac{\partial \mathbf{r}_i}{\partial q_j} \delta q_j$, where again we sum over repeated indicies (like j here). This means

$$\mathbf{F}_i \cdot \delta \mathbf{r}_i = \mathbf{F}_i \cdot \frac{\partial \mathbf{r}_i}{\partial q_j} \delta q_j \equiv Q_j \delta q_j \quad (1.65)$$

where we have defined *generalized forces*

$$Q_j \equiv \mathbf{F}_i \cdot \frac{\partial \mathbf{r}_i}{\partial q_j}. \quad (1.66)$$

We can also transform the $\dot{\mathbf{p}}_i \cdot \delta \mathbf{r}_i$ term using our earlier point transformation results as well as the fact that $\frac{d}{dt} \left(\frac{\partial \mathbf{r}_i}{\partial q_j} \right) = \frac{\partial^2 \mathbf{r}_i}{\partial q_j \partial t} + \sum_k \frac{\partial^2 \mathbf{r}_i}{\partial q_j \partial q_k} \dot{q}_k = \frac{\partial \mathbf{v}_i}{\partial q_j}$. Writing out the index sums explicitly,

this gives

$$\begin{aligned}
 \sum_i \dot{\mathbf{p}}_i \cdot \delta \mathbf{r}_i &= \sum_{i,j} m_i \ddot{\mathbf{r}}_i \cdot \frac{\partial \mathbf{r}_i}{\partial q_j} \delta q_j \\
 &= \sum_{i,j} \left(\frac{d}{dt} \left(m_i \dot{\mathbf{r}}_i \cdot \frac{\partial \mathbf{r}_i}{\partial q_j} \right) - m_i \dot{\mathbf{r}}_i \cdot \frac{d}{dt} \left(\frac{\partial \mathbf{r}_i}{\partial q_j} \right) \right) \delta q_j \\
 &= \sum_{i,j} \left(\frac{d}{dt} \left(m_i \mathbf{v}_i \cdot \frac{\partial \mathbf{v}_i}{\partial \dot{q}_j} \right) - m_i \mathbf{v}_i \cdot \frac{\partial \mathbf{v}_i}{\partial \dot{q}_j} \right) \delta q_j \\
 &= \sum_j \left(\frac{d}{dt} \left(\frac{\partial T}{\partial \dot{q}_j} \right) - \frac{\partial T}{\partial q_j} \right) \delta q_j
 \end{aligned} \tag{1.67}$$

for $T = \frac{1}{2} \sum_i m_i \mathbf{v}_i^2$. Together with the D'Alembert principle, we obtain the final result

$$\sum_j \left(\frac{d}{dt} \left(\frac{\partial T}{\partial \dot{q}_j} \right) - \frac{\partial T}{\partial q_j} - Q_j \right) \delta q_j = 0. \tag{1.68}$$

We will see momentarily that this result is somewhat more general than the Euler-Lagrange equations, containing them as a special case.

Holonomic Constraints:

We will start by considering systems with only holonomic constraints, postponing other types of constraints to the next section. Here we can find the independent coordinates q_j with $j = 1, \dots, N - k$ that satisfy the k constraints. This implies that the generalized virtual displacements δq_j are independent, so that their coefficients in Eq. (1.68) must vanish,

$$\frac{d}{dt} \left(\frac{\partial T}{\partial \dot{q}_j} \right) - \frac{\partial T}{\partial q_j} - Q_j = 0, \quad \text{for } j = 1, \dots, N - k. \tag{1.69}$$

There are several special cases of this result, which we derived from the d'Alembert principle.

1. For a conservative force $\mathbf{F}_i = -\nabla_i V$, then

$$Q_j = -(\nabla_i V) \cdot \frac{\partial \mathbf{r}_i}{\partial q_j} = -\frac{\partial V}{\partial q_j} \tag{1.70}$$

where we assume that the potential can be expressed in the generalized coordinates as $V = V(q, t)$. Then using $L = T - V$, we see that Eq. (1.69) simply reproduces the Euler-Lagrange equations $\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_j} \right) - \frac{\partial L}{\partial q_j} = 0$.

2. If $Q_j = -\frac{\partial V}{\partial q_j} + \frac{d}{dt} \left(\frac{\partial V}{\partial \dot{q}_j} \right)$ for $V = V(q, \dot{q}, t)$, which is the case for velocity dependent forces derivable from a potential (like the electromagnetic Lorentz force), then the Euler-Lagrange equations $\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_j} \right) - \frac{\partial L}{\partial q_j} = 0$ are again reproduced.

3. If Q_j has forces obtainable from a potential as in case 2, as well as generalized forces R_j that cannot, then

$$\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_j} \right) - \frac{\partial L}{\partial q_j} = R_j \quad (1.71)$$

is the generalization of the Euler-Lagrange equations with non-conservative generalized forces.

We will continue to use the notation R_j for generalized forces that can not be obtained from a potential. An important example of such nonconservative forces R_j is given by *friction*.

- Static friction is $F_s \leq F_s^{\max} = \mu_s F_N$ for a normal force F_N .
- Sliding friction is $\mathbf{F} = -\mu F_N \frac{\mathbf{v}}{v}$, so this is a constant force that is always opposite the direction of motion (but vanishes when there is no motion).
- Rolling friction is $\mathbf{F} = -\mu_R F_N \frac{\mathbf{v}}{v}$.
- Fluid friction at a low velocity is $\mathbf{F} = -bv \frac{\mathbf{v}}{v} = -b\mathbf{v}$.

A general form for a friction force is $\mathbf{F}_i = -h_i(v_i) \frac{\mathbf{v}_i}{v_i}$ (where as a reminder there is no implicit sum on i here since we specified i on the left-hand-side). For this form

$$R_j = - \sum_i h_i \frac{\mathbf{v}_i}{v_i} \cdot \frac{\partial \mathbf{r}_i}{\partial q_j} = - \sum_i h_i \frac{\mathbf{v}_i}{v_i} \cdot \frac{\partial \mathbf{v}_i}{\partial \dot{q}_j}. \quad (1.72)$$

Simplifying further gives

$$\begin{aligned} R_j &= - \sum_i \frac{h_i}{2v_i} \frac{\partial}{\partial \dot{q}_j} (v_i^2) = - \sum_i h_i \frac{\partial v_i}{\partial \dot{q}_j} = - \sum_i \frac{\partial v_i}{\partial \dot{q}_j} \frac{\partial}{\partial v_i} \int_0^{v_i} dv'_i h_i(v'_i) = - \frac{\partial}{\partial \dot{q}_j} \sum_i \int_0^{v_i} dv'_i h_i(v'_i) \\ &= - \frac{\partial \mathcal{F}}{\partial \dot{q}_j} \end{aligned} \quad (1.73)$$

where

$$\mathcal{F} = \sum_i \int_0^{v_i} dv'_i h_i(v'_i) \quad (1.74)$$

is the “dissipation function”. This is a scalar function like L so it is relatively easy to work with.

Example: Consider a sphere of radius a and mass m falling in a viscous fluid. Then $T = \frac{1}{2}m'\dot{y}^2$ where $m' < m$ accounts for the mass of displaced fluid (recall Archimedes principle that the buoyant force on a body is equal to the weight of fluid the body displaces). Also

$V = m'gy$, and $L = T - V$. Here $h \propto \dot{y}$, so $\mathcal{F} = 3\pi\eta a\dot{y}^2$, where by the constant of proportionality is determined by the constant η , which is the viscosity. From this, $\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{y}} \right) - \frac{\partial L}{\partial y} = -\frac{\partial \mathcal{F}}{\partial \dot{y}}$ gives the equation of motion $m'\ddot{y} + m'g = -6\pi\eta a\dot{y}$. The friction force $6\pi\eta a\dot{y}$ is known as Stokes Law. (We will derive this equation for the friction force from first principles later on, in our discussion of fluids.) This differential equation can be solved by adding a particular solution $y_p(t)$ to a solution of the homogeneous equation $m'\ddot{y}_H + 6\pi\eta a\dot{y}_H = 0$. For the time derivatives the results are $\dot{y}_p = -m'g/(6\pi\eta a)$ and $\dot{y}_H = A \exp(-6\pi\eta a t/m')$, where the constant A must be determined by an initial condition. The result $\dot{y} = \dot{y}_H + \dot{y}_p$ can be integrated in time once more to obtain the full solution $y(t)$ for the motion.

Example: if we add sliding friction to the case of two masses on a plane connected by a spring (considered on problem set #1), then $h_i = \mu_f m_i g$ for some friction coefficient μ_f , and

$$\mathcal{F} = \mu_f g(m_1 v_1 + m_2 v_2) = \mu_f g \left(m_1 \sqrt{\dot{x}_1^2 + \dot{y}_1^2} + m_2 \sqrt{\dot{x}_2^2 + \dot{y}_2^2} \right). \quad (1.75)$$

If we switch to a suitable set of generalized coordinates q_j that simplify the equations of motion without friction, and then compute the generalized friction forces $R_j = -\frac{\partial \mathcal{F}}{\partial \dot{q}_j}$, we can get the equations of motion including friction. Further details of how this friction complicates the equations of motion were provided by an example done in lecture.

1.5 Calculus of Variations & Lagrange Multipliers

Calculus of Variations

In the calculus of variations, we wish to find a set of functions $y_i(s)$ between s_1 and s_2 that extremize the following functional (a function of functions),

$$J[y_i] = \int_{s_1}^{s_2} ds f(y_1(s), \dots, y_n(s), \dot{y}_1(s), \dots, \dot{y}_n(s), s), \quad (1.76)$$

where for this general discussion only we let $\dot{y}_i \equiv \frac{dy_i}{ds}$ rather than $\frac{d}{dt}$. To consider the action of the functional under a variation we consider $y'_i(s) = y_i(s) + \eta_i(s)$ where $\eta_i(s_1) = \eta_i(s_2) = 0$, meaning that while the two endpoints are fixed during the variation $\delta y_i = \eta_i$, the path in between is varied. Expanding the variation of the functional integral $\delta J = J[y'_i] - J[y_i] = 0$ to 1st order in δy_i we have

$$0 = \delta J = \int_{s_1}^{s_2} ds \sum_i \left[\delta y_i \frac{\partial f}{\partial y_i} + \delta \dot{y}_i \frac{\partial f}{\partial \dot{y}_i} \right]. \quad (1.77)$$

Using integration by parts on the second term, and the vanishing of the variation at the endpoints to remove the surface term, δJ vanishes when

$$\int_{s_1}^{s_2} \sum_i \left[\frac{\partial f}{\partial y_i} - \frac{d}{ds} \left(\frac{\partial f}{\partial \dot{y}_i} \right) \right] \delta y_i(s) ds = 0. \quad (1.78)$$

If we identify a set of independent variations δy_j (for example, after imposing holonomic constraints), this can only occur if

$$\frac{\partial f}{\partial y_j} - \frac{d}{ds} \left(\frac{\partial f}{\partial \dot{y}_j} \right) = 0. \quad (1.79)$$

The scope of this calculus of variation result for extremizing the integral over f is more general than its application to classical mechanics.

Example: Hamilton's principle states that motion $q_i(t)$ extremizes the action, so in this case $s = t$, $y_i = q_i$, $f = L$, and $J = S$. Demanding $\delta S = 0$ then yields the Euler-Lagrange equations of motion from Eq. (1.79).

Example: As an example outside of classical mechanics, consider showing that the shortest distance between points on a sphere of radius a are great circles. This can be seen by minimizing the distance $J = \int_{s_1}^{s_2} ds$ where for a spherical surface,

$$ds = \sqrt{(dx)^2 + (dy)^2 + (dz)^2} = \sqrt{a^2(d\theta)^2 + a^2 \sin^2(\theta)(d\phi)^2} \quad (1.80)$$

since $dr = 0$. Taking $s = \theta$ and $y = \phi$, then

$$ds = a \sqrt{1 + \sin^2(\theta) \left(\frac{d\phi}{d\theta} \right)^2} d\theta, \quad (1.81)$$

so $f = \sqrt{1 + \sin^2(\theta) \dot{\phi}^2}$. The solution for the minimal path is given by solving $\frac{d}{d\theta} \left(\frac{\partial f}{\partial \dot{\phi}} \right) - \frac{\partial f}{\partial \phi} = 0$. After some algebra these are indeed found to be great circles, described by $\sin(\phi - \alpha) = \beta \cot(\theta)$ where α, β are constants.

Example: Hamilton's principle can also be used to yield the Hamilton equations of motion, by considering the variation of a path in phase space. In this case

$$\delta J[q, p] = \delta \int_{t_1}^{t_2} dt \left[p_i \dot{q}_i - H(q, p, t) \right] = 0 \quad (1.82)$$

must be solved with fixed endpoints: $\delta q_i(t_1) = \delta q_i(t_2) = 0$ and $\delta p_i(t_1) = \delta p_i(t_2) = 0$. Here, the role of y_i is played by the $2N$ variables $(q_1, \dots, q_N, p_1, \dots, p_N)$. As $f = p_i \dot{q}_i - H$, then

$$\begin{aligned} \frac{d}{dt} \left(\frac{\partial f}{\partial \dot{q}_i} \right) - \frac{\partial f}{\partial q_i} &= 0 & \implies & \dot{p}_i = -\frac{\partial H}{\partial q_i}, \\ \frac{d}{dt} \left(\frac{\partial f}{\partial \dot{p}_i} \right) - \frac{\partial f}{\partial p_i} &= 0 & \implies & \dot{q}_i = \frac{\partial H}{\partial p_i}, \end{aligned} \quad (1.83)$$

giving Hamilton's equations as expected. Note that because f is independent of \dot{p}_i , the term $(\partial f / \partial \dot{p}_i) \delta \dot{p}_i = 0$, and it would seem that we do not really need the condition that $\delta p_i(t_1) = \delta p_i(t_2) = 0$ to remove the surface term. However, these conditions on the variations δp_i are actually *required* in order to put q_i and p_i on the same footing, including making transformation that mix up these variables (we will exploit this in detail later on when discussing canonical transformations).

It is interesting and useful to note that D'Alembert's principle

$$\left(\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_j} \right) - \frac{\partial L}{\partial q_j} - R_j \right) \delta q_j = 0 \quad (1.84)$$

is a “differential” version of the equations that encode the classical dynamics, while Hamilton's principle

$$\delta J = \int_{t_1}^{t_2} dt \left(\frac{\partial L}{\partial q_j} - \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_j} \right) \right) \delta q_j = 0 \quad (1.85)$$

(for $R_j = 0$ where all forces come from a potential) is an integrated version.

Method of Lagrange Multipliers

Next we will consider the method of *Lagrange multipliers*. For simplicity we will assume there are no generalized forces outside the potential, $R_j = 0$, until further notice. (It is straightforward to add back the R_j to the final equations derived below.) The method of Lagrange multipliers will be useful for two situations that we will encounter:

1. When we actually want to study the forces of constraint that are holonomic.
2. When we have semi-holonomic constraints.

Let us consider k constraints for n coordinates, with $\alpha \in \{1, \dots, k\}$ being the index running over the constraints. These holonomic or semi-holonomic constraints take the form

$$g_\alpha(q, \dot{q}, t) = a_{j\alpha}(q, t) \dot{q}_j + a_{t\alpha}(q, t) = 0 \quad (1.86)$$

where again repeated indices are summed. Thus, $g_\alpha dt = a_{j\alpha} dq_j + a_{t\alpha} dt = 0$. For a virtual displacement δq_j we have $dt = 0$, so

$$\sum_{j=1}^n a_{j\alpha} \delta q_j = 0, \quad (1.87)$$

which gives us k equations constraining the virtual displacements. For each equation we can multiply by a function $\lambda_\alpha(t)$ known as *Lagrange multipliers*, and sum over α , and the combination will still be zero. Adding this zero to D'Alembert's principle (taking $R_j = 0$) yields

$$\left[\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_j} \right) - \frac{\partial L}{\partial q_j} - \lambda_\alpha a_{j\alpha} \right] \delta q_j = 0, \quad (1.88)$$

where the sums implicitly run over both α and j . It is clear that the Lagrange multiplier term is zero if we sum over j first, but now we want to consider summing first over α for a fixed j . Our goal is make the term in square brackets zero. Only $n - k$ of the virtual displacements δq_j are independent, so for these values of j the square brackets must vanish. For the remaining k values of j we can simply choose the k Lagrange multipliers λ_α to force the k square bracketed equations to be satisfied. This is known as the method of Lagrange multipliers. Thus all square bracketed terms are zero, and we have the generalization of the Euler-Lagrange equations which includes terms for the constraints:

$$\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_j} \right) - \frac{\partial L}{\partial q_j} = \lambda_\alpha a_{j\alpha}. \quad (1.89)$$

This is n equations, for the n possible values of j , and on the right-hand-side we sum over α for each one of these equations. The sum $\lambda_\alpha a_{j\alpha}$ can be interpreted as a generalized constraint force Q_j . The Lagrange multipliers λ_α and generalized coordinates q_j together form $n + k$ parameters, and equation (1.89) in conjunction with $g_\alpha = 0$ for each α from (1.86) together form $n + k$ equations to be solved.

There are two important cases to be considered.

1. In the holonomic case, $f_\alpha(q, t) = 0$. Here, $g_\alpha = \dot{f}_\alpha = \frac{\partial f_\alpha}{\partial q_j} \dot{q}_j + \frac{\partial f_\alpha}{\partial t}$, so $a_{j\alpha} = \frac{\partial f_\alpha}{\partial q_j}$. This gives

$$\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_j} \right) - \frac{\partial L}{\partial q_j} = \sum_{\alpha=1}^k \lambda_\alpha \frac{\partial f_\alpha}{\partial q_j} \quad (1.90)$$

for holonomic constraints. The same result can be derived from a generalized Hamilton's principle

$$J[q_j, \lambda_\alpha] = \int_{t_1}^{t_2} (L + \lambda_\alpha f_\alpha) dt \quad (1.91)$$

by demanding that $\delta J = 0$. It is convenient to think of $-\lambda_\alpha f_\alpha$ as an extra potential energy that we add into L so that a particle does work if it leaves the surface defined by $f_\alpha = 0$. Recall that given this potential, the $\text{Force}_q = -\nabla_q(-\lambda_\alpha f_\alpha) = \lambda_\alpha \nabla_q f_\alpha$, where the derivative $\nabla_q f_\alpha$ gives a vector that is normal to the constraint surface of constant $f_\alpha = 0$. This agrees with the form of our generalized force above.

2. In the semi-holonomic case, we just have $g_\alpha = a_{j\alpha}(q, t)\dot{q}_j + a_{t\alpha}(q, t) = 0$, with $a_{j\alpha} = \frac{\partial g_\alpha}{\partial \dot{q}_j}$. This gives

$$\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_j} \right) - \frac{\partial L}{\partial q_j} = \sum_{\alpha=1}^k \lambda_\alpha \frac{\partial g_\alpha}{\partial \dot{q}_j} \quad (1.92)$$

for semi-holonomic constraints. This result cannot be derived from Hamilton's principle in general. This justifies the time we spent discussing d'Alembert's principle, which we have used to obtain (1.92). Recall that static friction imposes a no-slip constraint in the form of our equation $g_\alpha = 0$. For $g \propto \dot{q}$, the form $\frac{\partial g}{\partial \dot{q}}$, is consistent with the form of generalized force we derived from our dissipation function, $\frac{\partial \mathcal{F}}{\partial \dot{q}}$ in our discussion of friction.

We end this chapter with several examples of the use of Lagrange multipliers.

Example: Consider a particle of mass m at rest on the top of a sphere of radius a , as shown above in Fig. 1.7. The particle is given an infinitesimal displacement $\theta = \theta_0$ so that it slides down. At what angle does it leave the sphere?

We use the coordinates (r, θ, ϕ) but set $\phi = 0$ by symmetry as it is not important. The constraint $r \geq a$ is non-holonomic, but while the particle is in contact with the sphere the constraint $f = r - a = 0$ is holonomic. To answer this question we will look for the point where the constraint force vanishes. Here $T = \frac{m}{2} (\dot{r}^2 + r^2 \dot{\theta}^2)$ and $V = mgz = mgr \cos(\theta)$ so that $L = T - V$, then $\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{r}} \right) - \frac{\partial L}{\partial r} = \lambda \frac{\partial f}{\partial r}$ gives

$$m\ddot{r} - mr\dot{\theta}^2 + mg \cos(\theta) = \lambda, \quad (1.93)$$

while $\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{\theta}} \right) - \frac{\partial L}{\partial \theta} = \lambda \frac{\partial f}{\partial \theta} = 0$ gives

$$\frac{d}{dt} (mr^2 \dot{\theta}) - mgr \sin(\theta) = 0. \quad (1.94)$$

This in conjunction with $r = a$ gives 3 equations for the 3 variables (r, θ, λ) . At this point we can now consider solving the equations together. The constraint equation gives $\dot{r} = 0$ so $\ddot{r} = 0$. This means

$$ma^2 \ddot{\theta} = mga \sin(\theta), \quad -ma\dot{\theta}^2 + mg \cos(\theta) = \lambda.$$

Multiply the first of these by $\dot{\theta}$ and integrate over time, knowing that $\dot{\theta} = 0$ when $\theta = 0$, gives $\dot{\theta}^2 = \frac{2g}{a} (1 - \cos(\theta))$. Plugging this into the second equation we find that

$$\lambda = mg(3 \cos(\theta) - 2) \quad (1.95)$$

is the radial constraint force. The mass leaves the sphere when $\lambda = 0$ which is when $\cos(\theta) = \frac{2}{3}$ (so $\theta \approx 48^\circ$).

What if we instead imposed the constraint $f' = r^2 - a^2 = 0$? If we call its Lagrange multiplier λ' we would get $\lambda' \frac{\partial f'}{\partial r} = 2a\lambda'$ when $r = a$, so $2a\lambda' = \lambda$ is the constraint force from before. The meaning of λ' is different, and it has different units, but we still have the same constraint force.

What are the equations of motion for $\theta > \arccos(\frac{2}{3})$? Now we no longer have the constraint so

$$m\ddot{r} - mr\dot{\theta}^2 + mg\cos(\theta) = 0 \quad \text{and} \quad \frac{d}{dt}(mr^2\dot{\theta}) - mgr\sin(\theta) = 0.$$

The initial conditions are $r_1 = a$, $\theta_1 = \arccos(\frac{2}{3})$, $\dot{r}_1 = 0$, and $\dot{\theta}_1^2 = \frac{2g}{3a}$ from before. Simpler coordinates are $x = r\sin(\theta)$ and $z = r\cos(\theta)$, giving

$$L = \frac{m}{2}(\dot{x}^2 + \dot{z}^2) - mgz, \quad (1.96)$$

so $\ddot{x} = 0$ and $\ddot{z} = -g$ with initial conditions $z_1 = \frac{2a}{3}$, $x_1 = \frac{\sqrt{5}a}{3}$, and the initial velocities $v_{z1} = \dot{z}(t_1) = -a\sin\theta_1\dot{\theta}_1 = -\frac{\sqrt{5}a}{3}\left(\frac{2g}{3a}\right)^{1/2}$ and $v_{x1} = \dot{x}(t_1) = a\cos\theta_1\dot{\theta}_1 = \frac{2a}{3}\left(\frac{2g}{3a}\right)^{1/2}$. Here t_1 is the time when the mass leaves the sphere. This means

$$x(t) = v_{x1}(t - t_1) + x_1, \quad (1.97)$$

$$z(t) = -\frac{g}{2}(t - t_1)^2 + v_{z1}(t - t_1) + z_1. \quad (1.98)$$

The time t_1 can be found from

$$\dot{\theta}^2 = \frac{2g}{a}(1 - \cos(\theta)) = \frac{4g}{a}\sin^2\left(\frac{\theta}{2}\right), \quad (1.99)$$

so $t_1 = \sqrt{\frac{a}{4g}} \int_{\theta_0}^{\arccos(\frac{2}{3})} \frac{d\theta}{\sin(\frac{\theta}{2})}$ where θ_0 is the small initial angular displacement from the top of the sphere. Note that the calculation of t_1 is the only place where we see that it was important to have a small θ_0 different from zero.

Example: Consider a hoop of radius a and mass m rolling down an inclined plane of angle ϕ without slipping as shown in Fig. 1.11, where we define the \hat{x} direction as being parallel to the ramp as shown. What is the friction force of constraint, and how does the acceleration compare to the case where the hoop is sliding rather than rolling?

The no-slip constraint means $a\dot{\theta} = \dot{x}$, so $h = a\dot{\theta} - \dot{x} = 0$, which can be made holonomic but which we will treat as semi-holonomic. Here the kinetic energy has contributions from the center mass motion of the hoop and its rotation, $T = T_{\text{CM}} + T_{\text{rotation}} = \frac{1}{2}m\dot{x}^2 + \frac{1}{2}ma^2\dot{\theta}^2$, where the hoops moment of inertia is $I_{\text{hoop}} = ma^2$. (We will discuss moments of inertia in detail in chapter 2.) Meanwhile, $V = mg(l - x)\sin(\phi)$ where we are free to pick the length l where $V(x = l) = 0$, such as at the bottom of the ramp. This means

$$L = T - V = \frac{m}{2}\dot{x}^2 + \frac{ma^2}{2}\dot{\theta}^2 + mg(x - l)\sin(\phi). \quad (1.100)$$

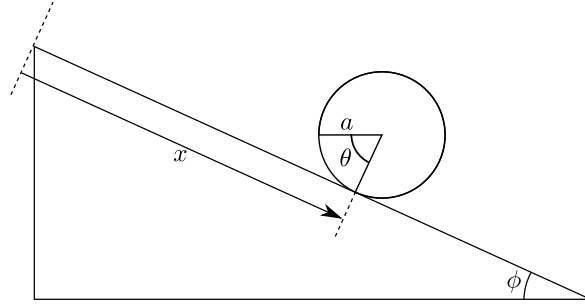


Figure 1.11: Hoop rolling on inclined plane

The equations of motion from $\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{x}} \right) - \frac{\partial L}{\partial x} = \lambda \frac{\partial h}{\partial x}$ and $\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{\theta}} \right) - \frac{\partial L}{\partial \theta} = \lambda \frac{\partial h}{\partial \theta}$ are

$$m\ddot{x} - mg \sin(\phi) = -\lambda \quad \text{and} \quad ma^2\ddot{\theta} = \lambda a, \quad (1.101)$$

along with $\dot{x} = a\dot{\theta}$. Taking a time derivative of the constraint gives $\ddot{x} = a\ddot{\theta}$, so $m\ddot{x} = \lambda$, and $\ddot{x} = \frac{g}{2} \sin(\phi)$. This is one-half of the acceleration of a sliding mass. Plugging this back in we find that

$$\lambda = \frac{1}{2}mg \sin(\phi) \quad (1.102)$$

is the friction force in the $-\hat{x}$ direction for the no-sliding constraint, and also $\ddot{\theta} = \frac{g}{2a} \sin(\phi)$.

Example: Consider a wedge of mass m_2 and angle α resting on ice and moving without friction. Let us also consider a mass m_1 sliding without friction on the wedge and try to find the equations of motion and constraint forces. The constraints are that $y_2 = 0$ so the

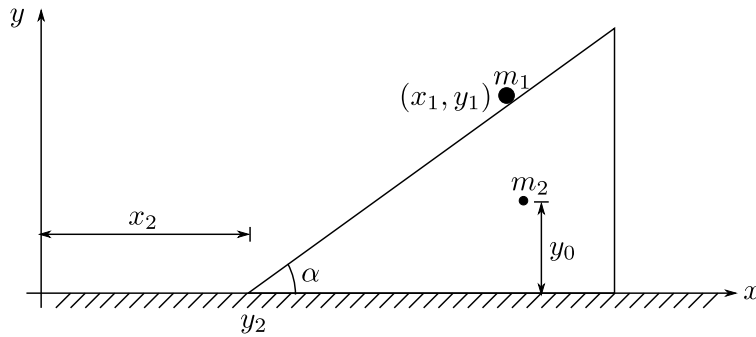


Figure 1.12: Wedge sliding on ice

wedge is always sitting on ice, and $\frac{y_1 - y_2}{x_1 - x_2} = \tan(\alpha)$ so the point mass is always sitting on the wedge. We will ignore the constraint force for no rotation of the wedge, and aim to find the constrain forces for these two constraints, and the equations for the motion.

The kinetic energy is simply $T = \frac{m_1}{2} (\dot{x}_1^2 + \dot{y}_1^2) + \frac{m_2}{2} (\dot{x}_2^2 + \dot{y}_2^2)$, while the potential energy is $V = m_1 g y_1 + m_2 g (y_2 + y_0)$, where y_0 is the height of the center of mass of the wedge. Then $L = T - V$, with the constraints $f_1 = (y_1 - y_2) - (x_1 - x_2) \tan(\alpha) = 0$ and $f_2 = y_2 = 0$. The equations of motion from the Euler-Lagrange equations with these holonomic constraints are

$$\begin{aligned} \frac{d}{dt} \frac{\partial L}{\partial \dot{x}_1} - \frac{\partial L}{\partial x_1} &= \lambda_1 \frac{\partial f_1}{\partial x_1} + \lambda_2 \frac{\partial f_2}{\partial x_1} &\implies m_1 \ddot{x}_1 &= -\lambda_1 \tan(\alpha), \\ \frac{d}{dt} \frac{\partial L}{\partial \dot{y}_1} - \frac{\partial L}{\partial y_1} &= \lambda_1 \frac{\partial f_1}{\partial y_1} + \lambda_2 \frac{\partial f_2}{\partial y_1} &\implies m_1 \ddot{y}_1 + m_1 g &= \lambda_1, \\ \frac{d}{dt} \frac{\partial L}{\partial \dot{x}_2} - \frac{\partial L}{\partial x_2} &= \lambda_1 \frac{\partial f_1}{\partial x_2} + \lambda_2 \frac{\partial f_2}{\partial x_2} &\implies m_2 \ddot{x}_2 &= \lambda_1 \tan(\alpha), \\ \frac{d}{dt} \frac{\partial L}{\partial \dot{y}_2} - \frac{\partial L}{\partial y_2} &= \lambda_1 \frac{\partial f_1}{\partial y_2} + \lambda_2 \frac{\partial f_2}{\partial y_2} &\implies m_2 \ddot{y}_2 + m_2 g &= -\lambda_1 + \lambda_2, \end{aligned} \quad (1.103)$$

which in conjunction with $y_1 - y_2 = (x_1 - x_2) \tan(\alpha)$ and $y_2 = 0$ is six equations for six variables. We number them (1) to (6) and now consider putting the equations together. Equation (6) gives $\ddot{y}_2 = 0$ so (4) gives $m_2 g = \lambda_2 - \lambda_1$ where λ_2 is the force of the ice on the wedge and λ_1 is the *vertical* force (component) of the wedge on the point mass. Adding (1) and (3) gives $m_1 \ddot{x}_1 + m_2 \ddot{x}_2 = 0$ meaning that the CM of m_1 and m_2 has no overall force acting on it.

Additionally, as (5) implies $\ddot{y}_1 = (\ddot{x}_1 - \ddot{x}_2) \tan(\alpha)$, then using (1), (2), and (3) we find the constant force

$$\lambda_1 = \frac{g}{\frac{1}{m_1 \cos^2(\alpha)} + \frac{\tan^2(\alpha)}{m_2}}. \quad (1.104)$$

With this result in hand we can use it in (1), (2), and (3) to solve for the trajectories. Since

$$\begin{aligned} \ddot{x}_2 &= \frac{\tan(\alpha)}{m_2} \lambda_1, \\ \ddot{x}_1 &= -\frac{\tan(\alpha)}{m_1} \lambda_1, \\ \ddot{y}_1 &= \frac{\lambda_1}{m_1} - g, \end{aligned} \quad (1.105)$$

the accelerations are constant. As a check on our results, if $m_2 \rightarrow \infty$, then $\ddot{x}_2 = 0$ so indeed the wedge is fixed; and for this case, $\ddot{x}_1 = -g \sin(\alpha) \cos(\alpha)$ and $\ddot{y}_1 = -g \sin^2(\alpha)$ which both vanish as $\alpha \rightarrow 0$ as expected (since in that limit the wedge disappears, flattening onto the icy floor below it).

Chapter 2

Rigid Body Dynamics

2.1 Coordinates of a Rigid Body

A set of N particles forms a *rigid body* if the distance between any 2 particles is fixed:

$$r_{ij} \equiv |\mathbf{r}_i - \mathbf{r}_j| = c_{ij} = \text{constant}. \quad (2.1)$$

Given these constraints, how many generalized coordinates are there?

If we know 3 non-collinear points in the body, the remaining points are fully determined by triangulation. The first point has 3 coordinates for translation in 3 dimensions. The second point has 2 coordinates for spherical rotation about the first point, as r_{12} is fixed. The third point has one coordinate for circular rotation about the axis of \mathbf{r}_{12} , as r_{13} and r_{23} are fixed. Hence, there are *6 independent coordinates*, as represented in Fig. 2.1. The result that there are six continuous coordinates is independent of N , so it also applies to a continuous body (in the limit of $N \rightarrow \infty$).

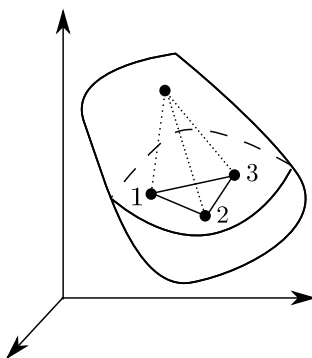


Figure 2.1: Three non-collinear points can be fully determined by using 6 continuous coordinates. Since the distances between any two other points are fixed in the rigid body, any other point of the body is fully determined by the distance to these 3 points (and a discrete orientation which is not a dynamical variable).

The translations of the body require three spatial coordinates. These translations can be taken from any fixed point in the body. Typically the fixed point is the center of mass (CM), defined as:

$$\mathbf{R} = \frac{1}{M} \sum_i m_i \mathbf{r}_i, \quad (2.2)$$

where m_i is the mass of the i -th particle and \mathbf{r}_i the position of that particle with respect to a fixed origin and set of axes (which will notationally be unprimed) as in Fig. 2.2. In the case of a continuous body, this definition generalizes as:

$$\mathbf{R} = \frac{1}{M} \int_{\mathcal{V}} \mathbf{r} \rho(\mathbf{r}) d\mathcal{V}, \quad (2.3)$$

where $\rho(\mathbf{r})$ is the mass density at position \mathbf{r} and we integrate over the volume \mathcal{V} .

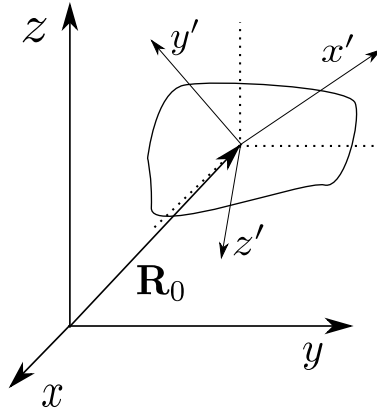


Figure 2.2: The three translational coordinates correspond to the position of the Center of Mass, and the three rotational coordinates correspond to the three angles necessary to define the orientation of the axis fixed with the body.

Rotations of the body are taken by fixing axes with respect to the body (we will denote these body fixed axes with primes) and describing their orientation with respect to the unprimed axes by 3 angles (ϕ, θ, ψ) .

A particularly useful choice of angles are called *Euler angles*. The angle ϕ is taken as a rotation about the z -axis, forming new \tilde{x} - and \tilde{y} -axes while leaving the z -axis unchanged, as shown in Fig. 2.3. The angle θ is then taken as a rotation about the \tilde{x} -axis, forming new \tilde{y}' - and z' -axes while leaving the \tilde{x} -axis unchanged, as shown in Fig. 2.4. Finally, the angle ψ is taken as a rotation about the z' -axis, forming new x' - and y' -axes while leaving the z' -axis unchanged, as shown in Fig. 2.5. (The \tilde{x} -axis is called the line of nodes, as it is the intersection of the xy - and $x'y'$ -planes.)

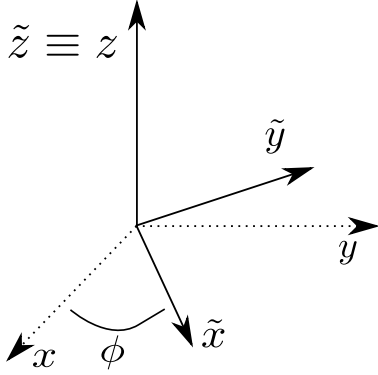


Figure 2.3: First rotation is by ϕ around the original z axis.

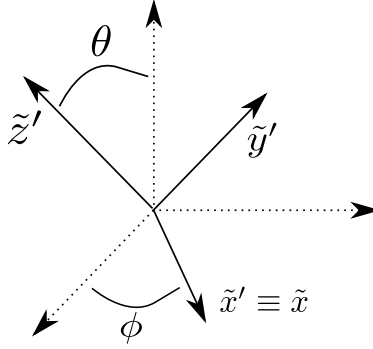


Figure 2.4: Second rotation is by θ around the intermediate \tilde{x} axis.

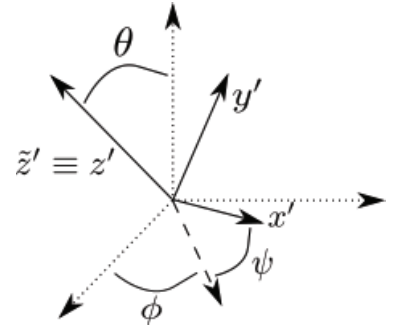


Figure 2.5: Final rotation is by ψ around the final z' axis.

Rotations can be described by 3×3 matrices U . This means each rotation step can be described as a matrix multiplication. Where $\mathbf{r} = (x, y, z)$, then

$$\tilde{\mathbf{r}} = U_{\phi} \mathbf{r} = \begin{bmatrix} \cos(\phi) & \sin(\phi) & 0 \\ -\sin(\phi) & \cos(\phi) & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}. \quad (2.4)$$

Similar transformations can be written for the other terms:

$$\tilde{\mathbf{r}}' = U_{\theta} \tilde{\mathbf{r}} \quad , \quad \mathbf{r}' = U_{\psi} \tilde{\mathbf{r}}' = U_{\psi} U_{\theta} \tilde{\mathbf{r}} = U_{\psi} U_{\theta} U_{\phi} \mathbf{r}.$$

Defining the total transformation as U , it can be written as:

$$U \equiv U_{\psi} U_{\theta} U_{\phi} \Rightarrow \mathbf{r}' = U \mathbf{r}. \quad (2.5)$$

Care is required with the order of the terms since the matrices don't commute. Writing U out explicitly:

$$U = \begin{bmatrix} \cos(\psi) & \sin(\psi) & 0 \\ -\sin(\psi) & \cos(\psi) & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos(\theta) & \sin(\theta) \\ 0 & -\sin(\theta) & \cos(\theta) \end{bmatrix} \begin{bmatrix} \cos(\phi) & \sin(\phi) & 0 \\ -\sin(\phi) & \cos(\phi) & 0 \\ 0 & 0 & 1 \end{bmatrix}. \quad (2.6)$$

All rotation matrices, including U_{ϕ} , U_{θ} , U_{ψ} , and U are *orthogonal*. Orthogonal matrices W satisfy

$$W^{\top} W = W W^{\top} = \mathbb{1} \quad \Leftrightarrow \quad W^{\top} = W^{-1}, \quad (2.7)$$

where $\mathbb{1}$ refers to the identity matrix and \top to the transpose. This ensures that the length of a vector is invariant under rotations:

$$\mathbf{r}^2 = \mathbf{r}^{\top} (W^{\top} W) \mathbf{r} = \mathbf{r}^2. \quad (2.8)$$

Orthogonal matrices W have 9 entries but need to fulfill 6 conditions from orthogonality, leaving only 3 free parameters, corresponding to the 3 angles necessary to determine the rotation.

We can also view $\mathbf{r}' = U\mathbf{r}$ as a transformation from the vector \mathbf{r} to the vector \mathbf{r}' in the same coordinate system. This is an active transformation, as opposed to the previous perspective which was a passive transformation.

Finally, note that inversions like

$$U = \begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix} \quad \text{or} \quad U = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix} \quad (2.9)$$

are not rotations. These have $\det(U) = -1$, so they can be forbidden by demanding that $\det(U) = 1$. All orthogonal matrices have $\det(W) = \pm 1$ because $\det(W^T W) = (\det(W))^2 = 1$. In the language of group theory, the restriction to $\det(W) = 1$ gives the special orthogonal group $SO(3)$ as opposed to simply $O(3)$, the orthogonal group. We disregard the $\det(U) = -1$ subset of transformations because it is impossible for the system to undergo these transformations without the distance between the particles jumping abruptly in the process (they are not continuously connected to the identity transformation), so they are not allowed mechanical motions of our rigid body.

Intuitively, we could rotate the coordinates (x, y, z) directly into the coordinates (x', y', z') by picking the right axis of rotation. In fact, the *Euler theorem* states that a general displacement of a rigid body with one point fixed is equivalent to a single rotation about *some axis* that runs through the fixed point. This theorem will be true if a general rotation U leaves some axis fixed, which is satisfied by

$$U\mathbf{r} = \mathbf{r} \quad (2.10)$$

for any point \mathbf{r} on this axis. This is an eigenvalue equation for U with eigenvalue 1. To better understand this, we need to use the linear algebra concepts of eigenvalue equations, which we now review.

Although the notion of an eigenvalue equation generally holds for linear operators, for now the discussion will be restricted to orthogonal rotation matrices U . The eigenvalue equation is

$$U\boldsymbol{\xi} = \lambda\boldsymbol{\xi}, \quad (2.11)$$

where $\boldsymbol{\xi}$ is an eigenvector and λ is the associated eigenvalue. Rewriting this as

$$(U - \lambda\mathbb{1})\boldsymbol{\xi} = 0 \quad (2.12)$$

requires that $\det(U - \lambda\mathbb{1}) = 0$, so that $U - \lambda\mathbb{1}$ is not invertible and the solution can be non-trivial, $\boldsymbol{\xi} \neq 0$. $\det(U - \lambda\mathbb{1}) = 0$ is a cubic equation in λ , which has 3 solutions, which are the eigenvalues λ_α for $\alpha \in \{1, 2, 3\}$. The associated eigenvectors are $\boldsymbol{\xi}^{(\alpha)}$ and satisfy

$$U\boldsymbol{\xi}^{(\alpha)} = \lambda_\alpha\boldsymbol{\xi}^{(\alpha)}, \quad (2.13)$$

where no implicit sum over repeated indices is taken. Forming a matrix from the resulting eigenvectors as columns:

$$X = \begin{bmatrix} \uparrow & \uparrow & \uparrow \\ \boldsymbol{\xi}^{(1)} & \boldsymbol{\xi}^{(2)} & \boldsymbol{\xi}^{(3)} \\ \downarrow & \downarrow & \downarrow \end{bmatrix} \quad (2.14)$$

then we can rewrite Eq.(2.13) as

$$UX = X \cdot \text{diag}(\lambda_1, \lambda_2, \lambda_3) \Rightarrow X^{-1}UX = \text{diag}(\lambda_1, \lambda_2, \lambda_3) \quad (2.15)$$

This means X diagonalizes U . Since U is orthogonal, the matrix X is unitary ($X^\dagger X = XX^\dagger = \mathbb{1}$). Note that $^\top$ indicates transposition whereas † indicates Hermitian conjugation (i.e. complex conjugation * combined with transposition $^\top$).

Next we note that since $\det(U) = 1$, then $\lambda_1 \lambda_2 \lambda_3 = 1$. Additionally, $|\lambda_\alpha|^2 = 1$ for each α because:

$$U\boldsymbol{\xi} = \lambda\boldsymbol{\xi} \Rightarrow \boldsymbol{\xi}^\dagger U^\top = \lambda^* \boldsymbol{\xi}^\dagger \Rightarrow \boldsymbol{\xi}^\dagger \boldsymbol{\xi} = \boldsymbol{\xi}^\dagger U^\top U \boldsymbol{\xi} = |\lambda|^2 \boldsymbol{\xi}^\dagger \boldsymbol{\xi}. \quad (2.16)$$

Finally, if λ is an eigenvalue, then so is λ^* , since we can take the complex conjugate of the original equation:

$$U\boldsymbol{\xi} = \lambda\boldsymbol{\xi} \Rightarrow U\boldsymbol{\xi}^* = \lambda^* \boldsymbol{\xi}^* \quad (2.17)$$

where $\boldsymbol{\xi}^*$ is still a column vector but with its elements undergoing complex conjugation with respect to $\boldsymbol{\xi}$. Without loss of generality, let us say for a rotation matrix U that $\lambda_2 = \lambda_3^*$. Then $1 = \lambda_1 |\lambda_2|^2 = \lambda_1$, so one of the eigenvalues is 1, giving Eq.(2.10), and thus proving Euler's Theorem. The associated eigenvector $\boldsymbol{\xi}^{(1)}$ to the eigenvalue $\lambda_1 = 1$ is the rotation axis, and if $\lambda_2 = \lambda_3^* = e^{i\Phi}$ then Φ is the rotation angle about that axis.

In fact, we can make good use of our analysis of Euler's theorem. Together the rotation axis and rotation angle can be used to define the instantaneous *angular velocity* $\boldsymbol{\omega}(t)$ such that:

$$|\boldsymbol{\omega}| = \frac{d\Phi}{dt} \quad \text{and} \quad \boldsymbol{\omega} \parallel \boldsymbol{\xi}^{(1)}. \quad (2.18)$$

The angular velocity will play an important role in our discussion of time dependence with rotating coordinates in the next section. If we consider several consecutive displacements of the rigid body, then each can have its own axis $\boldsymbol{\xi}^{(1)}$ and its own $\dot{\Phi}$, so $\boldsymbol{\omega}$ changes at each instance of time, and hence $\boldsymbol{\omega} = \boldsymbol{\omega}(t)$ (for the entire rigid body).

2.2 Time Evolution with Rotating Coordinates

Lets use unprimed axes (x, y, z) for the fixed (inertial) axes, with fixed basis vectors \mathbf{e}_i . We will also use primed axes (x', y', z') for the body axes with basis vectors \mathbf{e}'_i .

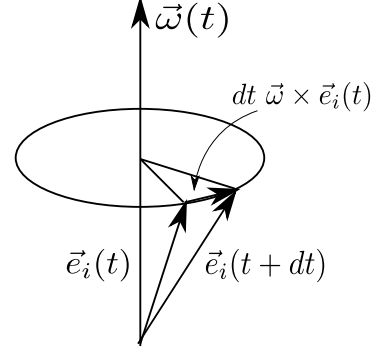
If we consider *any* vector then it can be decomposed with either set of basis vectors:

$$\mathbf{b} = \sum_i b_i \mathbf{e}_i = \sum_i b'_i \mathbf{e}'_i. \quad (2.19)$$

For fixed axes basis vectors by definition $\dot{\mathbf{e}}_i = 0$, while for those in the body frame,

$$\dot{\mathbf{e}}'_i = \boldsymbol{\omega}(t) \times \mathbf{e}'_i \quad (2.20)$$

meaning vectors of fixed length undergo a rotation at a time t . The derivation of this result is shown in the figure on the right, by considering the change to the vector after an infinitesimal time interval dt .



Summing over repeated indices, this means:

$$\dot{\mathbf{b}} = \dot{b}_i \mathbf{e}_i = \dot{b}'_i \mathbf{e}'_i + \boldsymbol{\omega}(t) \times (b'_i \mathbf{e}'_i) = \dot{b}'_i \mathbf{e}'_i + \boldsymbol{\omega}(t) \times \mathbf{b}$$

Defining $\frac{d}{dt}$ as the time evolution in the fixed (F) frame and $\frac{d_R}{dt}$ the time evolution in the rotating/body (R) frame, then vectors evolve in time according to

$$\frac{d\mathbf{b}}{dt} = \frac{d_R \mathbf{b}}{dt} + \boldsymbol{\omega} \times \mathbf{b}. \quad (2.21)$$

As a mnemonic we have the operation “ $(d/dt) = d_R/dt + \boldsymbol{\omega} \times$ ” which can act on any vector.

Let us apply this to the position \mathbf{r} of a particle of mass m , which gives

$$\frac{d\mathbf{r}}{dt} = \frac{d_R \mathbf{r}}{dt} + \boldsymbol{\omega} \times \mathbf{r} \quad \Leftrightarrow \quad \mathbf{v}_F = \mathbf{v}_R + \boldsymbol{\omega} \times \mathbf{r}. \quad (2.22)$$

Taking another time derivative gives us the analog for acceleration:

$$\begin{aligned} \frac{\mathbf{F}}{m} &= \frac{d\mathbf{v}_F}{dt} = \frac{d_R \mathbf{v}_F}{dt} + \boldsymbol{\omega} \times \mathbf{v}_F \\ &= \frac{d_R \mathbf{v}_R}{dt} + \frac{d_R \boldsymbol{\omega}}{dt} \times \mathbf{r} + \boldsymbol{\omega} \times \frac{d_R \mathbf{r}}{dt} + \boldsymbol{\omega} \times \mathbf{v}_R + \boldsymbol{\omega} \times (\boldsymbol{\omega} \times \mathbf{r}). \end{aligned} \quad (2.23)$$

As $\frac{d_R \mathbf{r}}{dt} = \mathbf{v}_R$ is the velocity within the body frame and $\frac{d_R \mathbf{v}_R}{dt} = \mathbf{a}_R$ is the acceleration within the body frame, and $d_R \boldsymbol{\omega}/dt = d\boldsymbol{\omega}/dt = \dot{\boldsymbol{\omega}}$ then

$$m\mathbf{a}_R = \mathbf{F} - m\boldsymbol{\omega} \times (\boldsymbol{\omega} \times \mathbf{r}) - 2m\boldsymbol{\omega} \times \mathbf{v}_R - m\dot{\boldsymbol{\omega}} \times \mathbf{r} \quad (2.24)$$

gives the acceleration in the body frame with respect to the forces that seem to be present in that frame. The terms $-m\boldsymbol{\omega} \times (\boldsymbol{\omega} \times \mathbf{r})$ and $-2m\boldsymbol{\omega} \times \mathbf{v}_R$ are, respectively, the centrifugal

and Coriolis fictitious forces respectively, while the last term $-m \frac{d_{\mathbf{R}} \boldsymbol{\omega}}{dt} \times \mathbf{r}$ is a fictitious force that arises from non-uniform rotational motion, so that there is angular acceleration within the body frame. The same result could also have been obtained with the Euler-Lagrange equations for L in the rotating coordinates:

$$L = \frac{m}{2} (\dot{\mathbf{r}} + \boldsymbol{\omega} \times \mathbf{r})^2 - V, \quad (2.25)$$

and you will explore this on a problem set.

Note that the centrifugal term is radially outward and perpendicular to the rotation axis. To see this, decompose \mathbf{r} into components parallel and perpendicular to $\boldsymbol{\omega}$, $\mathbf{r} = \mathbf{r}_{\parallel} + \mathbf{r}_{\perp}$, then $\boldsymbol{\omega} \times \mathbf{r} = \boldsymbol{\omega} \times \mathbf{r}_{\perp}$, so $-\boldsymbol{\omega} \times (\boldsymbol{\omega} \times \mathbf{r}_{\perp}) = \omega^2 \mathbf{r}_{\perp}$. This term is present for any rotating body. On the other hand, the Coriolis force is nonzero when there is a nonzero velocity in the rotating/body frame: $\mathbf{v}_{\mathbf{R}} \neq 0$.

Example: Consider the impact of the Coriolis force on projectile motion on the rotating Earth, where the angular velocity is $\omega_{\text{Earth}} = \frac{2\pi}{24 \times 3600 \text{ s}} \approx 7.3 \times 10^{-5} \text{ s}^{-1}$. We work out the cross-product $-\boldsymbol{\omega} \times \mathbf{v}_r$ as shown in Fig. 2.6 for a particle in the northern hemisphere, where ω points to the north pole. Thus a projectile in the northern/southern hemisphere would be perturbed to the right/left relative to its velocity direction \mathbf{v}_r .

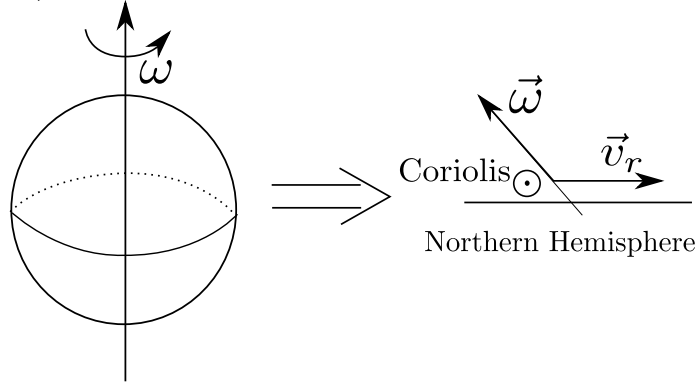
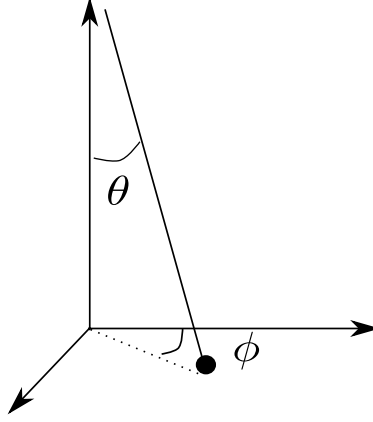
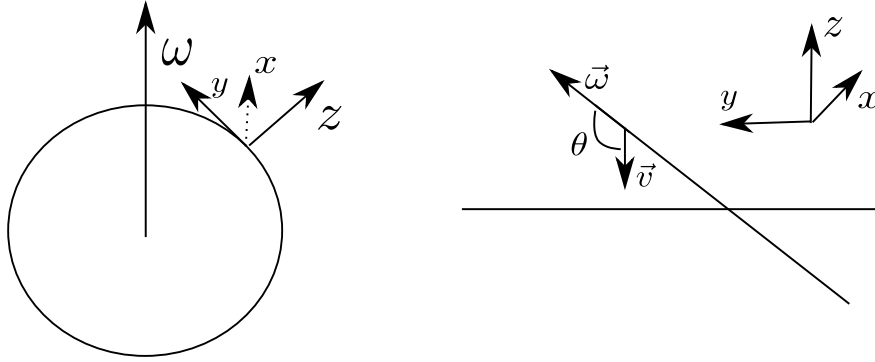


Figure 2.6: For a projectile, in the Northern Hemisphere, the Coriolis pushes it to its right, relative to its direction of motion.

Exercise: Consider a Foucault pendulum which hangs from a rigid rod, but is free to oscillate in two angular directions, as shown in Fig. 2.2. For $\theta \ll 1$ and working to first order in the small ω , the result derived from the Coriolis force gives $\dot{\phi} \approx \omega_{\text{Earth}} \sin(\lambda)$. Here λ is the latitude angle measured from equator. The precession is clockwise in the northern hemisphere, and is maximized at the north pole where $\lambda = 90^\circ$.



Example: Consider the Coriolis deflection of a freely falling body on Earth in the northern hemisphere. We use the coordinate system shown below, where z is perpendicular to the surface of the earth and y is parallel to the earth's surface and points towards the north pole.



Working to first order in the small ω gives us

$$m\mathbf{a}_R = m\dot{\mathbf{v}}_R = -mg\hat{z} - 2m\boldsymbol{\omega} \times \mathbf{v}, \quad (2.26)$$

where the centrifugal terms of order $O(\omega^2)$ are dropped. As an initial condition we take $\mathbf{v}(t=0) = v_0\hat{z}$. The term $-\boldsymbol{\omega} \times \mathbf{v}$ points along \hat{x} , so:

$$\ddot{z} = -g + O(\omega^2) \quad \Rightarrow \quad v_z = v_0 - gt \quad (2.27)$$

Moreover implementing the boundary condition that $\dot{x}(t=0) = 0$:

$$\ddot{x} = -2(\boldsymbol{\omega} \times \mathbf{v})_x = -2\omega \sin(\theta)v_z(t) \quad \Rightarrow \quad \dot{x} = -2\omega \sin(\theta) \left(v_0 t - \frac{g}{2} t^2 \right). \quad (2.28)$$

Taking also $x(t=0) = 0$, and integrating one further time, the motion in the x direction is:

$$x(t) = -2\omega \sin(\theta) \left(\frac{v_0}{2} t^2 - \frac{g}{6} t^3 \right). \quad (2.29)$$

Lets consider this effect for a couple simple cases. If the mass m is dropped from a height $z(t=0) = h_{\max}$ with zero velocity, $v_0 = 0$, then:

$$z = h_{\max} - \frac{g}{2}t^2 \quad (2.30)$$

and the mass reaches the floor at time

$$t_1 = \sqrt{\frac{2h_{\max}}{g}}. \quad (2.31)$$

From Eq.(2.28) we see that $\dot{x} > 0$ for all t , and that:

$$x(t = t_1) = \frac{2\sqrt{2}\omega \sin(\theta)h_{\max}^{3/2}}{3\sqrt{g}} > 0. \quad (2.32)$$

However, if the mass m is thrown up with an initial $\dot{z}(t=0) = v_0 > 0$ from the ground ($z = 0$), then :

$$z(t) = v_0 t - \frac{g}{2}t^2 > 0. \quad (2.33)$$

Here the particle rises to a maximum height $z = v_0^2/(2g)$ at time $t = v_0/g$, and then falls back to earth. Using Eq.(2.28) we see that $\dot{x} < 0$ for all t . If t_1 is the time it reaches the ground again ($t_1 = \frac{2v_0}{g}$), then:

$$x(t = t_1) = -\frac{4\omega \sin(\theta)v_0^3}{3g^2} < 0. \quad (2.34)$$

2.3 Kinetic Energy, Angular Momentum, and the Moment of Inertia Tensor for Rigid Bodies

Returning to rigid bodies, consider one built out of N fixed particles. The kinetic energy is best expressed using CM coordinates, where \mathbf{R} is the CM and we here take \mathbf{r}_i to be the *displacement* of particle i relative to the CM. Once again making sums over repeated subscripts as implicit, the kinetic energy (T) of the system is given by:

$$T = \frac{1}{2}M\dot{\mathbf{R}}^2 + \frac{1}{2}m_i\dot{\mathbf{r}}_i^2. \quad (2.35)$$

As the body is rigid, then points cannot translate relative to the body but can only rotate so that $\dot{\mathbf{r}}_i = \boldsymbol{\omega} \times \mathbf{r}_i$. The rotational kinetic energy is then

$$T_R = \frac{1}{2}m_i\dot{\mathbf{r}}_i^2 = \frac{1}{2}m_i(\boldsymbol{\omega} \times \mathbf{r}_i)^2 = \frac{1}{2}m_i[\boldsymbol{\omega}^2\mathbf{r}_i^2 - (\boldsymbol{\omega} \cdot \mathbf{r}_i)^2]. \quad (2.36)$$

Labeling Cartesian indices with a and b to reserve i and j for particle indices, then we can write out this result making the indices all explicit as

$$T_R = \frac{1}{2} \sum_{i,a,b} m_i (\delta_{ab} \mathbf{r}_i^2 - r_{ia} r_{ib}) \omega_a \omega_b. \quad (2.37)$$

It is convenient to separate out the parts in this formula that depend on the shape and distributions of masses in the body by defining the *moment of inertia tensor* \hat{I} for the discrete body as

$$\hat{I}_{ab} \equiv \sum_i m_i (\delta_{ab} \mathbf{r}_i^2 - r_{ia} r_{ib}). \quad (2.38)$$

The analog for a continuous body of mass density $\rho(\mathbf{r})$ is:

$$\hat{I}_{ab} \equiv \int_V (\mathbf{r}^2 \delta_{ab} - r_a r_b) \rho(\mathbf{r}) dV. \quad (2.39)$$

In terms of the moment of inertia tensor, the kinetic energy from rotation can now be written as:

$$T_R = \frac{1}{2} \sum_{a,b} \hat{I}_{ab} \omega_a \omega_b = \frac{1}{2} \boldsymbol{\omega} \cdot \hat{I} \cdot \boldsymbol{\omega}, \quad (2.40)$$

where in the last step we adopt a convenient matrix multiplication notation.

The moment of inertia tensor can be written with its components as a matrix in the form

$$\hat{I} = \sum_i m_i \begin{bmatrix} y_i^2 + z_i^2 & -x_i y_i & -x_i z_i \\ -x_i y_i & x_i^2 + z_i^2 & -y_i z_i \\ -x_i z_i & -y_i z_i & x_i^2 + y_i^2 \end{bmatrix}, \quad (2.41)$$

where the diagonal terms are the “moments of inertia” and the off-diagonal terms are the “products of inertia”. Note also that \hat{I} is symmetric in any basis, so $\hat{I}_{ab} = \hat{I}_{ba}$.

Special case: if the rotation happens about only one axis which can be defined as the z -axis for convenience so that $\boldsymbol{\omega} = (0, 0, \omega)$, then $T_R = \frac{1}{2} \hat{I}_{zz} \omega^2$ which reproduces the simpler and more familiar scalar form of the moment of inertia.

Lets now let \mathbf{r}_i be measured from a stationary point in the rigid body, which need not necessarily be the CM. The *angular momentum* can be calculated about this fixed point. Since $\mathbf{v}_i = \boldsymbol{\omega} \times \mathbf{r}_i$, we can write the angular momentum as:

$$\mathbf{L} = m_i \mathbf{r}_i \times \mathbf{v}_i = m_i \mathbf{r}_i \times (\boldsymbol{\omega} \times \mathbf{r}_i) = m_i [\mathbf{r}_i^2 \boldsymbol{\omega} - (\boldsymbol{\omega} \cdot \mathbf{r}_i) \mathbf{r}_i]. \quad (2.42)$$

Writing out the components

$$L_a = \sum_i m_i (\mathbf{r}_i^2 \omega_a - (\boldsymbol{\omega} \cdot \mathbf{r}_i) r_{ia}) = \sum_{i,b} \omega_b m_i (\delta_{ab} \mathbf{r}_i^2 - r_{ia} r_{ib}) = \sum_b \hat{I}_{ab} \omega_b, \quad (2.43)$$

which translates to the matrix equation:

$$\mathbf{L} = \hat{I} \cdot \boldsymbol{\omega}. \quad (2.44)$$

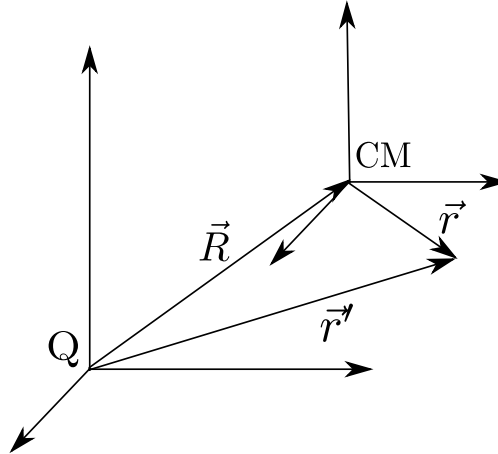
This allows us to write the corresponding rotational kinetic energy about the fixed point as:

$$T_R = \frac{1}{2} \boldsymbol{\omega} \cdot \mathbf{L}. \quad (2.45)$$

Note that in general, \mathbf{L} is *not* parallel to $\boldsymbol{\omega}$. We will see an explicit example of this below. Also note that the formula that we used for \hat{I} in this case is the same as we gave above. We use these formulas whether or not \mathbf{r}_i is taken with respect to the CM. (Do note that if the body is translating, such that no point in the body is fixed, then the CM point has a special role since the translational and rotational kinetic energy can be separated when using this point, and not in general other points.)

It is useful to pause to see what precisely the calculation of \hat{I} depends on. Since it involves components of the vectors r_i it depends on *the choice of the origin* for the rotation. Furthermore the entries of the matrix \hat{I}_{ab} obviously depend on the *orientation of the axes* used to define the components labeled by a and b . Given this, it is natural to ask whether given the result for \hat{I}_{ab} with one choice of axes and orientation, whether we can determine an $\hat{I}'_{a'b'}$ for a different origin and axes orientation. This is always possible with the help of a couple of theorems.

The parallel axis theorem: Given \hat{I}^{CM} about the CM, it is simple to find \hat{I}^{Q} about a different point Q with the *same* orientation for the axes. Referring to the figure below,



we define \mathbf{r}'_i as the coordinate of a particle i in the rigid body with respect to point Q and \mathbf{r}_i to be the coordinate of that particle with respect to the CM, so that:

$$\mathbf{r}'_i = \mathbf{R} + \mathbf{r}_i. \quad (2.46)$$

By definition of the CM:

$$\sum_i m_i \mathbf{r}_i = 0 \quad \text{and we let} \quad M = \sum_i m_i. \quad (2.47)$$

The tensor of inertia around the new axis is then:

$$\begin{aligned} \hat{I}_{ab}^Q &= m_i (\delta_{ab} \mathbf{r}_i'^2 - r_{ia}' r_{ib}') \\ &= m_i \left(\delta_{ab} (\mathbf{r}_i^2 + 2\mathbf{r}_i \cdot \mathbf{R} + \mathbf{R}^2) - \mathbf{r}_{ia} \mathbf{r}_{ib} - \mathbf{r}_{ia} \mathbf{R}_b - \mathbf{R}_a \mathbf{r}_{ib} - \mathbf{R}_a \mathbf{R}_b \right), \end{aligned} \quad (2.48)$$

where the cross terms involving a single \mathbf{r}_i or single component \mathbf{r}_{ia} sum up to zero by Eq.(2.47). The terms quadratic in \mathbf{r} are recognized as giving the moment of inertia tensor about the CM. This gives the *parallel axis theorem* for translating the origin:

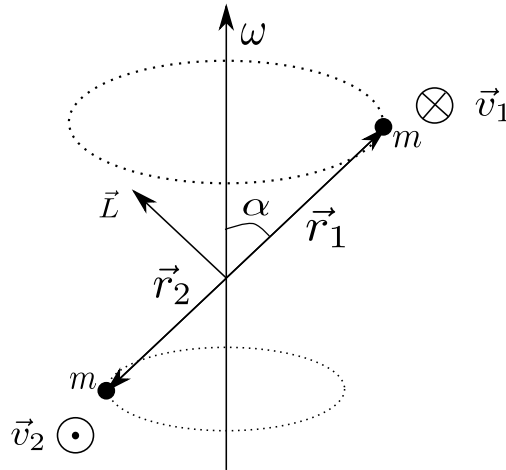
$$\hat{I}_{ab}^Q = M(\delta_{ab} \mathbf{R}^2 - \mathbf{R}_a \mathbf{R}_b) + \hat{I}_{ab}^{\text{CM}}, \quad (2.49)$$

If we wish to carry out a translation between P and Q , neither of which is the CM, then we can simply use this formula twice. Another formula can be obtained by projecting the parallel axis onto a specific axis \hat{n} where $\hat{n}^2 = 1$ (giving a result that may be familiar from an earlier classical mechanics course):

$$\begin{aligned} \hat{n} \cdot \hat{I}^Q \cdot \hat{n} &= M(\mathbf{R}^2 - (\hat{n} \cdot \mathbf{R})^2) + \hat{n} \cdot \hat{I}^{\text{CM}} \cdot \hat{n} = MR^2[1 - \cos^2(\theta)] + \hat{n} \cdot \hat{I}^{\text{CM}} \cdot \hat{n} \\ &= MR^2 \sin^2(\theta) + \hat{n} \cdot \hat{I}^{\text{CM}} \cdot \hat{n} \end{aligned} \quad (2.50)$$

where $\hat{n} \cdot \mathbf{R} \equiv R \cos(\theta)$.

Example: Lets consider an example of the calculation of \hat{I} for a situation where \mathbf{L} is not parallel to $\boldsymbol{\omega}$. Consider a dumbbell made of 2 identical point masses m attached by a massless rigid rod (but with different separations r_1 and r_2 from the axis of rotation), spinning so that $\boldsymbol{\omega} = \omega \hat{z}$ and so that the rod makes an angle α with the axis of rotation, as shown



We define body axes where the masses lie in the yz -plane. Here,

$$\mathbf{r}_1 = (0, r_1 \sin \alpha, r_1 \cos \alpha) \text{ and } \mathbf{r}_2 = (0, -r_2 \sin \alpha, -r_2 \cos \alpha). \quad (2.51)$$

Then using the definition of the moment inertia tensor:

$$\begin{aligned} I_{zz} &= m(x_1^2 + y_1^2) + m(x_2^2 + y_2^2) = m(r_1^2 + r_2^2) \sin^2 \alpha \\ I_{xx} &= m(y_1^2 + z_1^2) + m(y_2^2 + z_2^2) = m(r_1^2 + r_2^2) \\ I_{yy} &= m(x_1^2 + z_1^2) + m(x_2^2 + z_2^2) = m(r_1^2 + r_2^2) \cos^2 \alpha \\ I_{xy} &= I_{yx} = m(-x_1 y_1 - x_2 y_2) = 0 \\ I_{xz} &= I_{zx} = m(-x_1 z_1 - x_2 z_2) = 0 \\ I_{yz} &= I_{zy} = m(-y_1 z_1 - y_2 z_2) = -m(r_1^2 + r_2^2) \sin \alpha \cos \alpha \end{aligned} \quad (2.52)$$

Plugging these into $\mathbf{L} = \hat{I} \cdot \boldsymbol{\omega}$, recalling that only ω_z is non-zero, this gives

$$\mathbf{L} = (0, I_{yz}\omega, I_{zz}\omega). \quad (2.53)$$

Thus in this example \mathbf{L} is not parallel to $\boldsymbol{\omega}$.

Next, instead of translating the axes in a parallel manner, let us keep the origin fixed and rotate the axes according to an orthogonal rotation matrix U satisfying $U^\top U = U U^\top = 1$. Vectors are rotated as

$$\mathbf{L}' = U\mathbf{L} \quad , \quad \boldsymbol{\omega}' = U\boldsymbol{\omega} \quad \text{and therefore} \quad \boldsymbol{\omega} = U^\top \boldsymbol{\omega}'. \quad (2.54)$$

Putting these together

$$\mathbf{L}' = U\hat{I} \cdot \boldsymbol{\omega} = (U\hat{I}U^\top) \cdot \boldsymbol{\omega}' \quad \Rightarrow \quad \hat{I}' = U\hat{I}U^\top, \quad (2.55)$$

where \hat{I}' is the new moment of inertia tensor. (The fact that it transforms this way *defines* it as a tensor.) This allows us to calculate the new moment of inertia tensor after a rotation.

For a real symmetric tensor \hat{I} , there always exists a rotation from an orthogonal matrix U that diagonalizes \hat{I} giving a diagonal matrix \hat{I}' :

$$\hat{I}_D = \begin{pmatrix} I_1 & 0 & 0 \\ 0 & I_2 & 0 \\ 0 & 0 & I_3 \end{pmatrix}. \quad (2.56)$$

The entries of the diagonal moment of inertia tensor, I_α , are real and positive. This is just a special case of saying a Hermitian matrix can always be diagonalized by a unitary transformation (which is often derived in a Quantum Mechanics course as part of showing that a Hermitian matrix has real eigenvalues and orthogonal eigenvectors). The positivity

of diagonal matrix follows immediately from the definition of the moment of inertia tensor for the situation with zero off-diagonal terms.

The axes that make \hat{I} diagonal are called the *principal axes* and the components I_α are the *principal moments of inertia*. We find them by solving the eigenvalue problem

$$\hat{I} \cdot \boldsymbol{\xi} = \lambda \boldsymbol{\xi}, \quad (2.57)$$

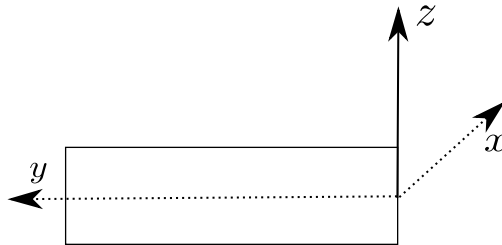
where the 3 eigenvalues λ give the principal moments of inertia I_α , and are obtained from solving $\det(\hat{I} - \lambda \mathbf{1}) = 0$. The corresponding 3 real *orthogonal* eigenvectors $\boldsymbol{\xi}^{(\alpha)}$ are the principal axes. Here $U^\top = [\boldsymbol{\xi}^{(1)} \quad \boldsymbol{\xi}^{(2)} \quad \boldsymbol{\xi}^{(3)}]$, where the eigenvectors fill out the columns. Then, without summing over repeated indices:

$$L_\alpha = I_\alpha \omega_\alpha \quad \text{and} \quad T = \frac{1}{2} \sum_\alpha I_\alpha \omega_\alpha^2, \quad (2.58)$$

where L_α and ω_α are the components of \mathbf{L} and $\boldsymbol{\omega}$, respectively, evaluated along the principal axes.

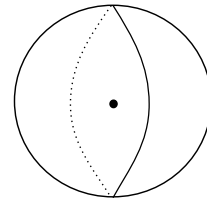
To summarize, for any choice of origin for any rigid body, there is a choice of axes that diagonalizes \hat{I} . For T to separate into translational and rotational parts, we must pick the origin to be the CM. Often, the principal axes can be identified by a symmetry of the body.

Example: for a thin rectangle lying in the yz -plane with one edge coinciding with the z -axis, and the origin chosen as shown below, then $I_{yz} = 0$ as the body is symmetric under $z \leftrightarrow -z$, while $I_{xz} = I_{xy} = 0$ as the body lies entirely within $x = 0$. Hence these are principal axes.

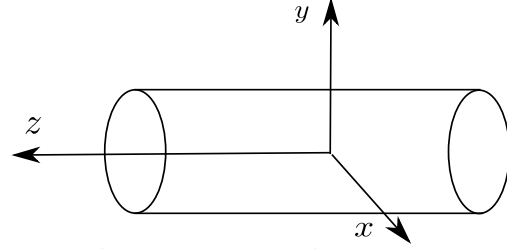


Sometimes, symmetry allows multiple choices for the principal axes.

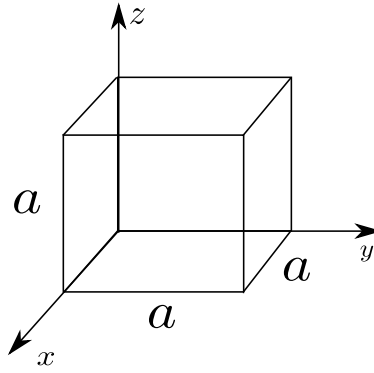
Example: for a sphere, any orthogonal axes through the origin are principal axes.



Example: for a cylinder whose central axis is aligned along the z -axis, because of rotational symmetry any choice of the x - and y -axes gives principal axes.



Example: Lets consider an example where the principal axes may not be apparent, which we can solve through the eigenvalue problem. Consider a uniform cube with sides of length a , mass m , and having the origin at one corner, as shown below.



By symmetry we have

$$\begin{aligned} I_{xx} = I_{yy} = I_{zz} &= \frac{m}{a^3} \int_0^a \int_0^a \int_0^a (x^2 + y^2) dx dy dz = \frac{2}{3}ma^2, \\ I_{xy} = I_{yz} = I_{xz} &= \frac{m}{a^3} \int_0^a \int_0^a \int_0^a -xz dx dy dz = -\frac{1}{4}ma^2. \end{aligned} \quad (2.59)$$

Thus the matrix is

$$\hat{I} = ma^2 \begin{pmatrix} +\frac{2}{3} & -\frac{1}{4} & -\frac{1}{4} \\ -\frac{1}{4} & +\frac{2}{3} & -\frac{1}{4} \\ -\frac{1}{4} & -\frac{1}{4} & +\frac{2}{3} \end{pmatrix}. \quad (2.60)$$

The principal moments of inertia are found from

$$\det(\hat{I} - \lambda \mathbf{1}) = \left(\frac{11}{12}ma^2 - \lambda \right)^2 \left(\frac{1}{6}ma^2 - \lambda \right) = 0. \quad (2.61)$$

This gives $I_1 = \lambda_1 = \frac{1}{6}ma^2$. Solving

$$(\hat{I} - \lambda_1 \mathbf{1})\boldsymbol{\xi}^{(1)} = 0 \quad \text{we find} \quad \boldsymbol{\xi}^{(1)} = (1, 1, 1). \quad (2.62)$$

The remaining eigenvalues are degenerate:

$$I_2 = I_3 = \lambda_2 = \lambda_3 = \frac{11}{12}ma^2 \quad (2.63)$$

so there is some freedom in determining the corresponding principal axes from $(\hat{I} - \lambda_2 \mathbb{1})\boldsymbol{\xi}^{(2,3)} = 0$, though they still should be orthogonal to each other (and $\boldsymbol{\xi}^{(1)}$). One example of a solution is:

$$\boldsymbol{\xi}^{(2)} = (1, -1, 0) \quad \text{and} \quad \boldsymbol{\xi}^{(3)} = (1, 1, -2) \quad (2.64)$$

Using these principal axes and the same origin, the moment of inertia tensor becomes

$$\hat{I}_D = \frac{ma^2}{6} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{11}{2} & 0 \\ 0 & 0 & \frac{11}{2} \end{pmatrix}. \quad (2.65)$$

In contrast, if we had chosen the origin as the center of the cube, then one choice for the principal axes would have the same orientation, but with $\hat{I}_{\text{CM}} = \frac{1}{6}ma^2 \mathbb{1}$. This result could be obtained from Eq. (2.65) using the parallel axis theorem. (Note that for a moment of inertia tensor $\propto \mathbb{1}$ that any set of orthogonal axes are principal axes due to the symmetry.)

2.4 Euler Equations

Consider the rotational motion of a rigid body about a fixed point (which could be the CM but could also be another point). We aim to describe the motion of this rigid body by exploiting properties of the body frame. To simplify things as much as possible, for this fixed point, we choose the principal axes fixed in the body frame indexed by $\alpha \in \{1, 2, 3\}$. Using the relation between time derivatives in the inertial and rotating frames, the torque is then given by:

$$\boldsymbol{\tau} = \frac{d\mathbf{L}}{dt} = \frac{d_R \mathbf{L}}{dt} + \boldsymbol{\omega} \times \mathbf{L} \quad (2.66)$$

where $\boldsymbol{\omega} = \boldsymbol{\omega}(t)$. For example:

$$\tau_1 = \frac{d_R L_1}{dt} + \omega_2 L_3 - \omega_3 L_2. \quad (2.67)$$

Not summing over repeated indices and using the formula for angular momentum along the principal axes gives $L_\alpha = I_\alpha \omega_\alpha$. Since we have fixed moments of inertia within the body we have $d_R I_\alpha / dt = 0$. Note that $d\boldsymbol{\omega}/dt = d_R \boldsymbol{\omega}/dt + \boldsymbol{\omega} \times \boldsymbol{\omega} = d_R \boldsymbol{\omega}/dt$, so its rotating and inertial time derivatives are the same, and we can write $\dot{\omega}_\alpha$ without possible cause of confusion. Thus $d_R L_\alpha / dt = I_\alpha \dot{\omega}_\alpha$. This yields *the Euler equations*:

$$\begin{aligned} I_1 \dot{\omega}_1 - (I_2 - I_3) \omega_2 \omega_3 &= \tau_1 \\ I_2 \dot{\omega}_2 - (I_3 - I_1) \omega_3 \omega_1 &= \tau_2 \\ I_3 \dot{\omega}_3 - (I_1 - I_2) \omega_1 \omega_2 &= \tau_3 \end{aligned} \quad (2.68)$$

where in all of these $\boldsymbol{\omega}$ and $\boldsymbol{\tau}$ are calculated in the rotating/body frame. This can also be written as

$$\tau_\alpha = I_\alpha \dot{\omega}_\alpha + \epsilon_{\alpha lk} \omega_l \omega_k I_k, \quad (2.69)$$

with α fixed but a sum implied over the repeated l and k indices. Here ϵ_{abc} is the fully antisymmetric Levi-Civita symbol.

Solving these equations gives $\omega_\alpha(t)$. Since the result is expressed in the body frame, rather than the inertial frame of the observer, this solution for $\boldsymbol{\omega}(t)$ may not always make the physical motion transparent. To fix this we can connect our solution to the Euler angles using the relations

$$\begin{aligned} \omega_1 = \omega_{x'} &= \dot{\phi} \sin \theta \sin \psi + \dot{\theta} \cos \psi, \\ \omega_2 = \omega_{y'} &= \dot{\phi} \sin \theta \cos \psi - \dot{\theta} \sin \psi, \\ \omega_3 = \omega_{z'} &= \dot{\phi} \cos \theta + \dot{\psi}. \end{aligned} \quad (2.70)$$

These results should be derived as exercise for the student.

Example: let us consider the stability of rigid-body free rotations ($\boldsymbol{\tau} = 0$). Is a rotation $\boldsymbol{\omega} = \omega_1 \mathbf{e}_1$ about the principal axis \mathbf{e}_1 stable?

Perturbations can be expressed by taking $\boldsymbol{\omega} = \omega_1 \mathbf{e}_1 + \kappa_2 \mathbf{e}_2 + \kappa_3 \mathbf{e}_3$, where κ_2 and κ_3 are small and can be treated to 1st order. The Euler equations are:

$$\dot{\omega}_1 = \frac{(I_2 - I_3)}{I_1} \kappa_2 \kappa_3 = O(\kappa^2) \approx 0, \quad (2.71)$$

so ω_1 is constant at this order, and

$$\dot{\kappa}_2 = \frac{(I_3 - I_1)}{I_2} \omega_1 \kappa_3 \quad \text{and} \quad \dot{\kappa}_3 = \frac{(I_1 - I_2)}{I_3} \omega_1 \kappa_2. \quad (2.72)$$

Combining these two equations yields

$$\ddot{\kappa}_2 = \left[\frac{(I_3 - I_1)(I_1 - I_2)\omega_1^2}{I_2 I_3} \right] \kappa_2, \quad (2.73)$$

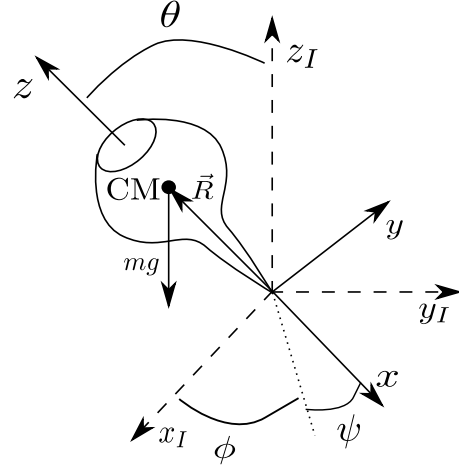
and the same equation for κ_3 . The term in the square bracket is constant, and is either negative $= -\omega^2$ with an oscillating solution $\kappa_2 \propto \cos(\omega t + \phi)$, or is positive $= \alpha^2$ with exponential solutions $\kappa_2 \propto a e^{\alpha t} + b e^{-\alpha t}$. If $I_1 < I_{2,3}$ or $I_{2,3} < I_1$ then the constant prefactor is negative, yielding stable oscillatory solutions. If instead $I_2 < I_1 < I_3$ or $I_3 < I_1 < I_2$ then the constant prefactor is positive, yielding an unstable exponentially growing component to their solution! This behavior can be demonstrated by spinning almost any object that has three distinct principal moments of inertia. A smart phone or tennis racket are good examples.

2.5 Symmetric Top with One Point Fixed

This section is devoted to a detailed analysis of a particular example that appears in many situations, the symmetric top with one point fixed, acted upon by a constant force.

Labeling the body axes as (x, y, z) and the fixed axes as (x_I, y_I, z_I) , as depicted in the right, symmetry implies that $I_1 = I_2$, and we will assume that $I_{1,2} \neq I_3$. The Euler angles are as usual (ϕ, θ, ψ) . From the figure we see that $\dot{\psi}$ is the rotation rate of the top about the (body) z -axis, $\dot{\phi}$ is the precession rate about the z_I fixed inertial axis, and $\dot{\theta}$ is the nutation rate by which the top may move away or towards the z_I axis. The Euler equations in this case are

$$\begin{aligned} I_1 \dot{\omega}_1 - (I_2 - I_3) \omega_2 \omega_3 &= \tau_1, \\ I_1 \dot{\omega}_2 - (I_3 - I_1) \omega_3 \omega_1 &= \tau_2, \\ I_3 \dot{\omega}_3 &= 0 = \tau_3. \end{aligned} \quad (2.74)$$



Since the CM coordinate \mathbf{R} is aligned along the z -axis there is no torque along z , $\tau_3 = 0$, leading to a constant ω_3 .

There are two main cases that we will consider.

Case: $\boldsymbol{\tau} = 0$ and $\dot{\theta} = 0$

The first case we will consider is when $\boldsymbol{\tau} = 0$ (so there is no gravity) and $\dot{\theta} = 0$ (so there is no nutation). Then

$$\frac{d\mathbf{L}}{dt} = \boldsymbol{\tau} = 0 \quad \Rightarrow \quad \mathbf{L} = \text{constant} \quad (2.75)$$

Let us define the constant:

$$\Omega \equiv \frac{I_3 - I_1}{I_1} \omega_3. \quad (2.76)$$

Then the Euler equations for this situation reduce to:

$$\dot{\omega}_1 + \Omega \omega_2 = 0 \quad \text{and} \quad \dot{\omega}_2 - \Omega \omega_1 = 0. \quad (2.77)$$

The simplest solution correspond to $\omega_1(t) = \omega_2(t) = 0$, where we just have a rotation about the z -axis. Here:

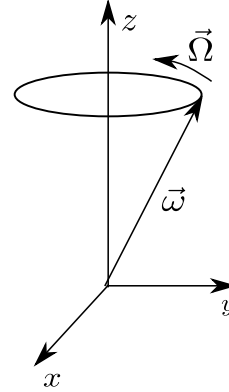
$$\begin{aligned} \mathbf{L} &= L_3 \mathbf{e}_3 \quad \text{where} \quad L_3 = I_3 \omega_3 \\ \omega_1 = \omega_2 = 0 &\quad \Rightarrow \quad \dot{\theta} = \dot{\phi} = 0 \quad \text{and} \quad \dot{\psi} = \omega_3. \end{aligned} \quad (2.78)$$

In this case $\mathbf{L} \parallel \boldsymbol{\omega}$. A more general situation is when \mathbf{L} and $\boldsymbol{\omega}$ are not necessarily parallel, and ω_1 and ω_2 do not vanish. In this case Eq. (2.77) is solved by:

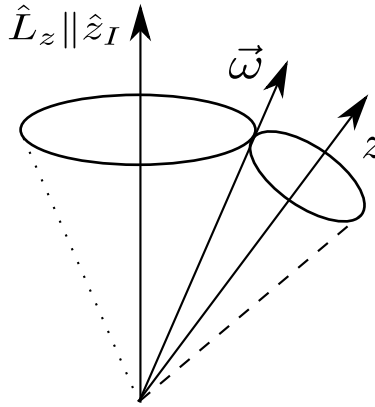
$$\omega_1 = C \sin(\Omega t + D) \quad \text{and} \quad \omega_2 = -C \cos(\Omega t + D). \quad (2.79)$$

The simple case corresponds to $C = 0$, so now we take $C > 0$ (since a sign can be accounted for by the constant phase D). This solution means $\boldsymbol{\omega}$ precesses about the body z -axis at the rate Ω , as pictured on the right. Since $\omega_1^2 + \omega_2^2$ is constant, the full $\omega = |\boldsymbol{\omega}|$ is constant, and is given by $\omega^2 = C^2 + \omega_3^2$.

The total energy here is just rotational kinetic energy $T_R = \frac{1}{2} \boldsymbol{\omega} \cdot \mathbf{L}$ which is constant too. Thus $\boldsymbol{\omega}$ also precesses about \mathbf{L} .



We can picture this motion by thinking about a body cone that rolls around a cone in the fixed coordinate system, where in the case pictured with a larger cone about \mathbf{L} we have $I_1 = I_2 > I_3$.



To obtain more explicit results for the motion we can relate Eq.(2.79) to Euler angles. Since $\dot{\theta} = 0$, we take $\theta = \theta_0$ to be constant. The other Euler angles come from:

$$\boldsymbol{\omega} = \begin{bmatrix} C \sin(\Omega t + D) \\ -C \cos(\Omega t + D) \\ \omega_3 \end{bmatrix} = \begin{bmatrix} \sin(\theta_0) \sin(\psi) \dot{\phi} \\ \sin(\theta_0) \cos(\psi) \dot{\phi} \\ \cos(\theta_0) \dot{\phi} + \dot{\psi} \end{bmatrix}. \quad (2.80)$$

Adding the squares of the 1st and 2nd components gives

$$C^2 = \sin^2(\theta_0) \dot{\phi}^2. \quad (2.81)$$

To be definite, take the positive square root of this equation to give

$$\dot{\phi} = \frac{C}{\sin(\theta_0)} \Rightarrow \phi = \frac{C}{\sin(\theta_0)} t + \phi_0. \quad (2.82)$$

The first two equations in Eq. (2.80) are then fully solved by taking $\psi = \pi - \Omega t - D$, so we find that both ϕ and ψ have linear dependence on time. Finally the third equation gives a relation between various constants

$$\omega_3 = C \cot(\theta_0) - \Omega. \quad (2.83)$$

Thus, we see that the solution has $\dot{\phi}$ and $\dot{\psi}$ are constants with $\dot{\theta} = 0$. If we had picked the opposite sign when solving Eq. (2.81) then we would have found similar results:

$$\dot{\phi} = -\frac{C}{\sin(\theta_0)} \Rightarrow \psi = -\Omega t - D \quad \text{and} \quad \omega_3 = -C \cot(\theta_0) - \Omega. \quad (2.84)$$

Case: $\tau \neq 0$ and $\dot{\theta} \neq 0$

Now we consider the general case where $\tau \neq 0$ and $\dot{\theta} \neq 0$. It is now more convenient to use the Lagrangian than the Euler equations directly. Since $I_1 = I_2$, using

$$T = \frac{1}{2} (I_1(\omega_1^2 + \omega_2^2) + I_3\omega_3^2) \quad \text{and} \quad \boldsymbol{\omega} = \begin{bmatrix} \sin(\theta) \sin(\psi) \dot{\phi} + \cos(\psi) \dot{\theta} \\ \sin(\theta) \cos(\psi) \dot{\phi} - \sin(\psi) \dot{\theta} \\ \cos(\theta) \dot{\phi} + \dot{\psi} \end{bmatrix}, \quad (2.85)$$

gives us the kinetic energy

$$T = \frac{I_1}{2} (\dot{\theta}^2 + \sin^2 \theta \dot{\phi}^2) + \frac{I_3}{2} (\dot{\psi} + \cos \theta \dot{\phi})^2. \quad (2.86)$$

Moreover, $V = mgR \cos(\theta)$, so in the Lagrangian $L = T - V$ the variables ϕ and ψ are cyclic. This means that the momenta

$$p_\phi = \frac{\partial L}{\partial \dot{\phi}} = [I_1 \sin^2(\theta) + I_3 \cos^2(\theta)] \dot{\phi} + I_3 \cos(\theta) \dot{\psi} \quad (2.87)$$

$$p_\psi = \frac{\partial L}{\partial \dot{\psi}} = I_3(\dot{\psi} + \cos(\theta) \dot{\phi}) = I_3 \omega_3 \quad (2.88)$$

are conserved (constant). Here p_ψ is same as the angular momentum L_3 discussed in the case above. The torque is along the line of nodes, and p_ϕ and p_ψ correspond to two projections of \mathbf{L} that are perpendicular to this torque (i.e. along \hat{z}_1 and \hat{z}). Additionally, the energy is given by

$$E = T + V = \frac{I_1}{2} (\dot{\theta}^2 + \sin^2(\theta) \dot{\phi}^2) + \frac{I_3}{2} (\dot{\psi} + \cos(\theta) \dot{\phi})^2 + mgR \cos(\theta) \quad (2.89)$$

and is also conserved. Solving the momentum equations, Eq. (2.87), for $\dot{\phi}$ and $\dot{\psi}$ gives

$$\begin{aligned}\dot{\phi} &= \frac{p_\phi - p_\psi \cos(\theta)}{I_1 \sin^2(\theta)} \\ \dot{\psi} &= \frac{p_\psi}{I_3} - \frac{(p_\phi - p_\psi \cos(\theta)) \cos(\theta)}{I_1 \sin^2(\theta)}.\end{aligned}\tag{2.90}$$

Note that once we have a solution for $\theta(t)$ that these two equations then allow us to immediately obtain solutions for $\phi(t)$ and $\psi(t)$ by integration. Eq. (2.90) can be plugged into the energy formula to give

$$E = \frac{I_1}{2} \dot{\theta}^2 + \frac{(p_\phi - p_\psi \cos(\theta))^2}{2I_1 \sin^2(\theta)} + \frac{p_\psi^2}{2I_3} + mgR \cos(\theta),\tag{2.91}$$

which is a (nonlinear) differential equation for θ , since all other quantities that appear are simply constants. To simplify this result take $u = \cos(\theta)$ so that:

$$1 - u^2 = \sin^2(\theta), \quad \dot{u} = -\sin(\theta)\dot{\theta}, \quad \dot{\theta}^2 = \frac{\dot{u}^2}{1 - u^2}.\tag{2.92}$$

Putting all this together gives:

$$\frac{\dot{u}^2}{2} = \left(\frac{2EI_3 - p_\psi^2}{2I_1 I_3} - \frac{mgR}{I_1} u \right) (1 - u^2) - \frac{1}{2} \left(\frac{p_\phi - p_\psi u}{I_1} \right)^2 \equiv -V_{\text{eff}}(u),\tag{2.93}$$

which is a cubic polynomial that we've defined to be the effective potential $V_{\text{eff}}(u)$. The solution to this from

$$dt = \pm \frac{du}{\sqrt{-2V_{\text{eff}}(u)}}\tag{2.94}$$

yields a complicated elliptic function, from which it is hard to get intuition for the motion.

Instead, we can look at the form of $V_{\text{eff}}(u)$, because

$$\frac{1}{2} \dot{u}^2 + V_{\text{eff}}(u) = 0\tag{2.95}$$

is the equation for the energy of a particle of unit mass $m = 1$, kinetic energy $\dot{u}^2/2$, a potential $V_{\text{eff}}(u)$, and with vanishing total energy. The cubic equation will have in general three roots where $V_{\text{eff}}(u) = 0$. Since the kinetic energy is always positive or zero, the potential energy must be negative or zero in the physical region, and hence the particle can not pass through any of the roots. The roots therefore serve as turning points. Furthermore, physical solutions are bounded by $-1 \leq (u = \cos \theta) \leq 1$. While the precise values for the roots will depend on the initial conditions or values of E , p_ψ , and p_ϕ , we can still describe the solutions in a fairly generic manner.

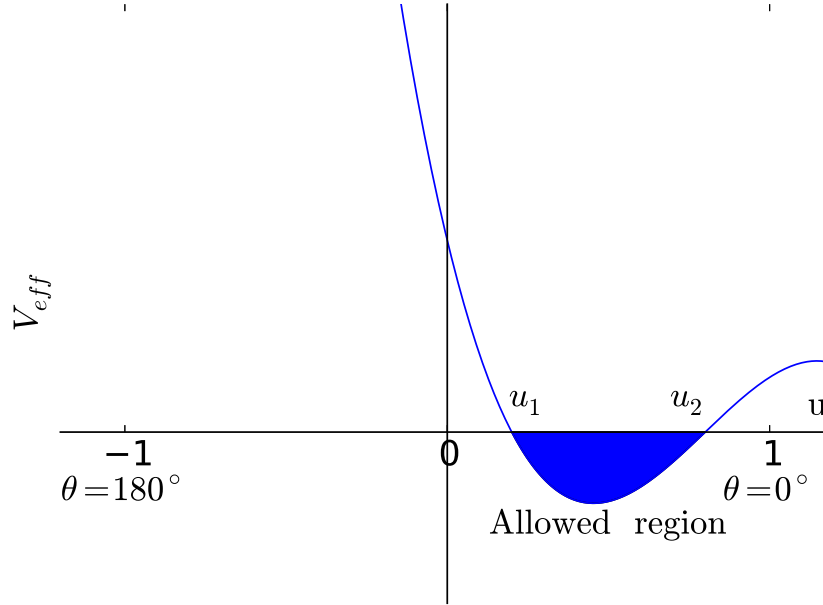


Figure 2.7: Allowed region for solutions for the top's nutation angle θ that solve Eq. (2.95).

Consider two roots u_1 and u_2 (corresponding respectively to some angles θ_1 and θ_2 as $u = \cos(\theta)$) satisfying $V_{\text{eff}}(u_1) = V_{\text{eff}}(u_2) = 0$, where $V_{\text{eff}}(u) < 0$ for $u_1 < u < u_2$; as shown in Fig. 2.7. We see that u_1 and u_2 correspond to the turning points of the motion. The region $u_1 < u < u_2$ corresponds to the region where the motion of our top lives and gives rise to a periodic nutation, where the solution bounces between the two turning points. Depending on the precise value of the various constants that appear in this V_{eff} this gives rise to different qualitative motions, with examples shown in Figs. 2.8–2.11. Recalling that $\dot{\phi} = (p_\phi - p_\psi u)/[I_1(1 - u^2)]$, we see that the possible signs for $\dot{\phi}$ will depend on p_ϕ and p_ψ . In Fig. 2.8 the top nutates between θ_1 and θ_2 while always precessing in the same direction with $\dot{\phi} > 0$, whereas in Fig. 2.9 the precession is also in the backward direction, $\dot{\phi} < 0$, for part of the range of motion. In Fig. 2.10 the top has $\dot{\phi} = 0$ at θ_2 , before falling back down in the potential and gaining $\dot{\phi} > 0$ again. This figure also captures the case where we let go of a top at $\theta = \theta_2 \geq 0$ that initially has $\dot{\psi} > 0$ but $\dot{\phi} = 0$. Finally in Fig. 2.11 we have the situation where there is no nutation oscillation because the two angles coincide, $\theta_1 = \theta_2$.

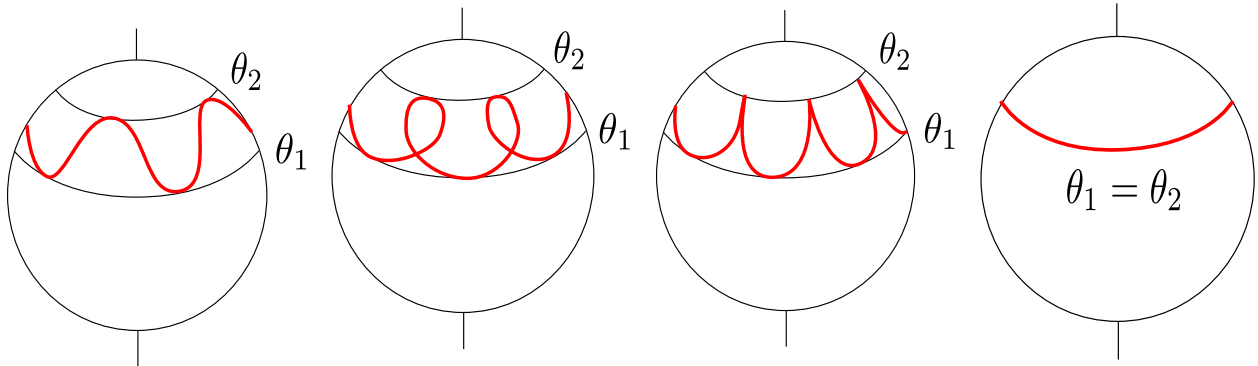


Figure 2.8: $\dot{\phi} > 0$

Figure 2.9: $\dot{\phi}$ has both signs

Figure 2.10: at θ_2 we have $\dot{\phi} = 0, \dot{\theta} = 0$

Figure 2.11: No nutation