Advanced Classical Mechanics

Prof. Iain W. Stewart

Classical Mechanics III
MIT 8.09 & 8.309

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

Last Compiled: October 6, 2021 – 20:19

Contents

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 $i \in \{1, \ldots, N\}$, or with generalized coordinates $q_i(t)$ for $i \in \{1, \ldots, 3N\}$ 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 2^{nd}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 $2N$ 1^{st} 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.

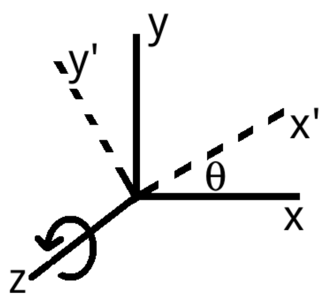


Figure 1.1: Frame rotated by an angle θ

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

$$\frac{d\hat{r}}{dt} = \dot{\theta}\hat{\theta}, \qquad \frac{d\hat{\theta}}{dt} = -\dot{\theta}\hat{r} \tag{1.1}$$

and thus

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

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

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

Here, $q = (q_1, \ldots, q_N)$ and likewise $\dot{q} = (\dot{q}_1, \ldots, \dot{q}_N)$. 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*,

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

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

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

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

   $$Q_i = Q_i(q_1, \ldots, q_N, t). \tag{1.6}$$

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

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

   we claim that the equations of motion are simply

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

   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

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

Therefore

$$\frac{\partial \dot{Q}}{\partial \dot{q}} = \frac{\partial Q}{\partial q}, \tag{1.10}$$

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

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

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

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

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,

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

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).

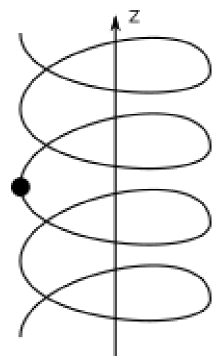


Figure 1.2: Bead on a helix

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 $q_2 = \theta$.

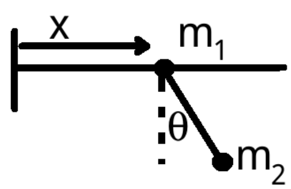


Figure 1.3: Pendulum of mass m_2 hanging on a rigid bar of length ℓ whose support m_1 is a frictionless horizontally sliding bead

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 $T = \frac{m}{2}\dot{\mathbf{r}}^2 = \frac{m}{2}\big(\dot{r}^2 + r^2\dot{\theta}^2\big)$, giving

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

For r the Euler-Lagrange equation is

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

This gives

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

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

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

This gives

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

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 $\tau = \mathbf{r} \times \mathbf{F}$.)

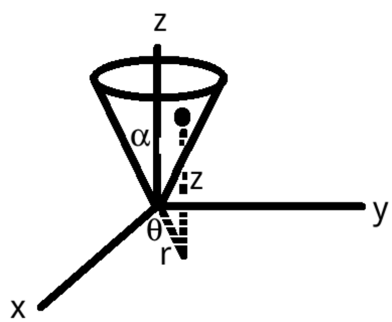


Figure 1.4: Particle on the inside of a cone

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 $T = \frac{m}{2}\left(\dot{r}^2 + r^2\dot{\theta}^2 + \dot{z}^2\right)$ 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

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

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

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

giving

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

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

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

giving

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

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

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

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

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

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

$$\frac{\partial H}{\partial t} = -\frac{\partial L}{\partial t}\,. \tag{1.26}$$

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

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

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):

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

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 equation have an have 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:

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

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 $\Longrightarrow$). 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

$$\dot{q}_i p_i = \dot{q}_i \frac{\partial L}{\partial \dot{q}_i} = \frac{1}{2}\dot{q}_i Tik \dot{q}_k + \frac{1}{2}\dot{q}_j T_{ji}\dot{q}_i = 2T. \tag{1.30}$$

Hence,

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

which is just the energy.

Another Special case: Consider a class of Lagrangians given as

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

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

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

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,

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

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 postive definite matrix. Thus, $H = \dot{\vec{q}} \cdot \vec{p} - L$ yields

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

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

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

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

$$y_1 = 0 \,, \qquad\qquad (y_1 - y_2)^2 + (x_1 - x_2)^2 = \ell^2 \,. \tag{1.37}$$

Prior to imposing any constraints the Lagrangian is

$$L = T - V = \frac{m_1}{2}\dot{x}_1^2 + \frac{m_2}{2}(\dot{x}_2^2 + \dot{y}_2^2) - m_2 g y_2 - m_1 g y_1\,. \tag{1.38}$$

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

$$L = \frac{m_1}{2}\dot{x}^2 + \frac{m_2}{2}\big(\dot{x}^2 + 2\ell\cos\theta\,\dot{x}\dot{\theta} + \ell^2\cos^2\theta\,\dot{\theta}^2 + \ell^2\sin^2\theta\,\dot{\theta}^2\big) + m_2 g\ell\cos\theta\,. \tag{1.39}$$

Next we determine the Hamiltonian. First we find

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

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

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

with $L = \frac{1}{2}(\dot{x}\;\dot{\theta})\cdot\hat{T}\cdot\begin{pmatrix} \dot{x} \\ \dot{\theta} \end{pmatrix} + L_0$ where $L_0 = m_2 g\ell\cos\theta$. Computing

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

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

$$\begin{aligned} H &= \frac{1}{2}(p_x\;p_\theta)\cdot\hat{T}^{-1}\cdot\begin{pmatrix} p_x \\ p_\theta \end{pmatrix} - m_2 g\ell\cos\theta \\ &= \frac{1}{2m_2\ell^2(m_1+m_2\sin^2\theta)}\Big[m_2\ell^2 p_x^2 + (m_1+m_2)p_\theta^2 - 2m_2\ell\cos\theta p_x p_\theta\Big] - m_2 g\ell\cos\theta\,. \end{aligned} \tag{1.43}$$

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

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

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

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

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

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

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

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

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 $\mathbf{R} = \frac{1}{M} m_i \mathbf{r}_i$, where the total mass $M = \sum_i m_i$. Then the CM momentum $\mathbf{P}$ is conserved. This is because

$$\mathbf{F}_{\mathbf{R}} = -\nabla_{\mathbf{R}} V = 0 \tag{1.48}$$

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

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

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 $\mathbf{F}_{i\to j} = -\mathbf{F}_{j\to i}$. Hence,

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

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

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

Simplifying further using the results in Fig. 2.2 yields

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

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,

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

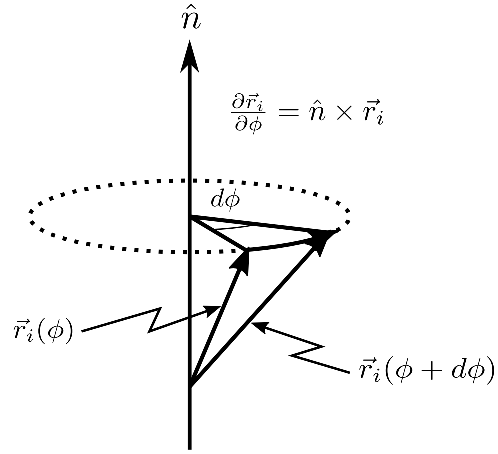


Figure 1.5: Rotation about a symmetry axis

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

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

3. One can also consider a scaling transformation. Suppose that under the transformation $\mathbf{r}_i \to \lambda \mathbf{r}_i$ the potential is homogeneous and transforms as $V \to \lambda^k V$ for some constant k. Letting T be quadratic in $\dot{\mathbf{r}}_i$ and taking time to transform as $t \to \lambda^{1-k/2}\, t$ then gives $\dot{\mathbf{r}}_i \to \lambda^{k/2}\, \dot{\mathbf{r}}_i$. So by construction $T \to \lambda^k T$ also, and thus the full Lagrangian $L \to \lambda^k L$. 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

$$\dot{H} = \frac{\partial H}{\partial t} = -\frac{\partial L}{\partial t} \,. \tag{1.55}$$

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

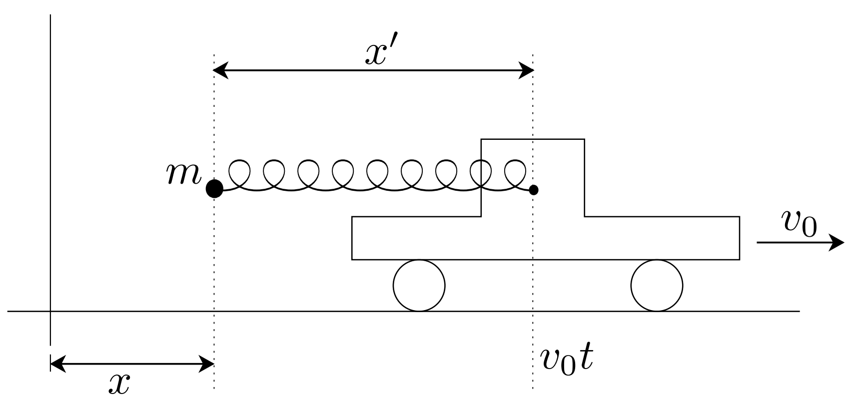


Figure 1.6: Mass attached by a spring to a moving cart

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

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

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

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

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

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

Note that $p' = m\dot{x}' + m v_0 = m\dot{x} = p$. This Lagrangian fits the general form in equation (1.32) with $a = m v_0$ and $L_0 = m v_0^2/2 - k x'^2/2$. So

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

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

$$f(q_1, \ldots, q_N, t) = 0 \tag{1.60}$$

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*.

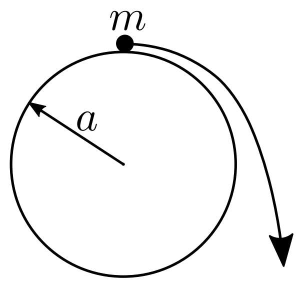


Figure 1.7: Mass on a sphere

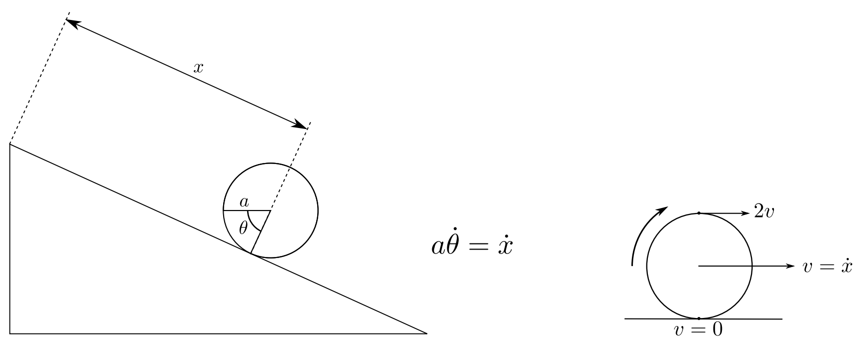


Figure 1.8: Disk rolling down an incline without slipping

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

$$\sum_j a_j(q)dq_j + a_t(q)dt = 0 \tag{1.61}$$

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

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

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

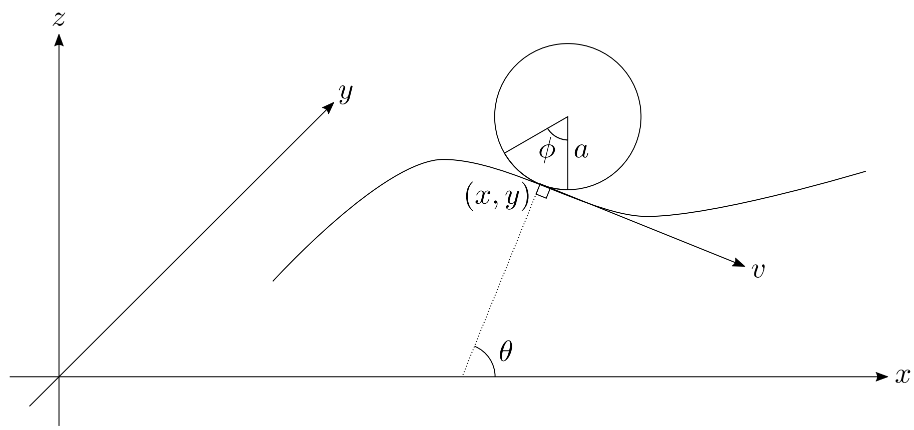


Figure 1.9: Vertical rolling disk on a two dimensional plane

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

$$\frac{\partial a_j}{\partial q_i} = \frac{\partial a_i}{\partial q_j}, \qquad \qquad \frac{\partial a_t}{\partial q_i} = \frac{\partial a_i}{\partial t}. \tag{1.63}$$

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 constrants 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 $x = a\varphi + x_0$, $y = y_0$, 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

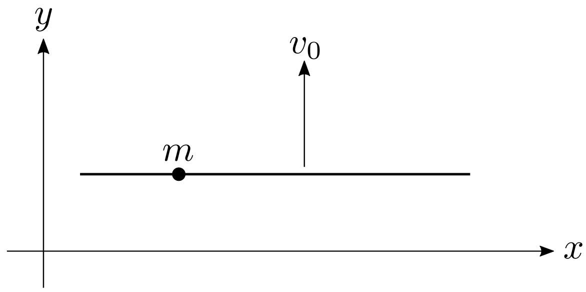


Figure 1.10: Bead on a moving wire

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

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

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

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

where we have defined *generalized forces*

$$Q_j \equiv \mathbf{F}_i\cdot\frac{\partial\mathbf{r}_i}{\partial q_j}\,. \tag{1.66}$$

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

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

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 $j = 1, \ldots, N-k$ 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,

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

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
$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_j}\right) - \frac{\partial L}{\partial q_j} = R_j \tag{1.71}$$
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 $F_{\mathrm{s}} \leq F_{\mathrm{s}}^{\max} = \mu_s F_{\mathrm{N}}$ for a normal force F_{N}.
- Sliding friction is $\mathbf{F} = -\mu F_{\mathrm{N}}\,\frac{\mathbf{v}}{v}$, so this is a constant force that is always opposite the direction of motion (but vanishes when there is no motion).
- Rolling friction is $\mathbf{F} = -\mu_{\mathrm{R}} F_{\mathrm{N}}\,\frac{\mathbf{v}}{v}$.
- 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
$$R_j = -\sum_i h_i \frac{\mathbf{v}_i}{v_i} \cdot \frac{\partial \mathbf{r}_i}{\partial q_j} = -\sum_i h_i \frac{\mathbf{v}_i}{v_i} \cdot \frac{\partial \mathbf{v}_i}{\partial \dot{q}_j}. \tag{1.72}$$

Simplifying further gives
$$\begin{aligned} R_j &= -\sum_i \frac{h_i}{2v_i}\frac{\partial}{\partial \dot{q}_j}\left(\mathbf{v}_i^2\right) = -\sum_i h_i \frac{\partial v_i}{\partial \dot{q}_j} = -\sum_i \frac{\partial v_i}{\partial \dot{q}_j}\frac{\partial}{\partial v_i}\int_0^{v_i} dv_i' h_i(v_i') = -\frac{\partial}{\partial \dot{q}_j}\sum_i \int_0^{v_i} dv_i' h_i(v_i') \\ &= -\frac{\partial \mathcal{F}}{\partial \dot{q}_j} \end{aligned} \tag{1.73}$$

where
$$\mathcal{F} = \sum_i \int_0^{v_i} dv_i'\, h_i(v_i') \tag{1.74}$$
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 $\dot{y}_H = A\exp(-6\pi\eta at/m')$, 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

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

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),

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

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 1$^{\text{st}}$ order in δy_i we have

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

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

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

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

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

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,

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

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

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

so $f = \sqrt{1 + \sin^2(\theta)\dot{\phi}^2}$. The solution for the minimal path is given by solving $\frac{d}{d\theta}\left(\frac{\partial f}{\partial \dot{\varphi}}\right) - \frac{\partial f}{\partial \varphi} = 0$. 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

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

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 $(q_1, \ldots, q_N, p_1, \ldots, p_N)$. As $f = p_i\dot{q}_i - H$, then

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

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

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

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

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

(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 $\alpha \in \{1, \ldots, k\}$ being the index running over the constraints. These holonomic or semi-holonomic constraints take the form

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

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

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

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

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

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:

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

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

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

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

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

   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
$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_j}\right) - \frac{\partial L}{\partial q_j} = \sum_{\alpha=1}^{k} \lambda_\alpha \frac{\partial g_\alpha}{\partial \dot{q}_j} \tag{1.92}$$
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 $T = \frac{m}{2}\left(\dot{r}^2 + r^2\dot{\theta}^2\right)$ 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
$$m\ddot{r} - mr\dot{\theta}^2 + mg\cos(\theta) = \lambda, \tag{1.93}$$
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
$$\frac{d}{dt}\left(mr^2\dot{\theta}\right) - mgr\sin(\theta) = 0. \tag{1.94}$$
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
$$ma^2\ddot{\theta} = mga\sin(\theta), \qquad -ma\dot{\theta}^2 + mg\cos(\theta) = \lambda.$$
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
$$\lambda = mg(3\cos(\theta) - 2) \tag{1.95}$$
is the radial constraint force. The mass leaves the sphere when $\lambda = 0$ which is when $\cos(\theta) = \frac{2}{3}$ (so $\theta \approx 48^o$).

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 $\theta > \arccos\left(\frac{2}{3}\right)$? Now we no longer have the constraint so

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

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

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

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

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

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

The time t_1 can be found from

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

so $t_1 = \sqrt{\frac{a}{4g}}\int_{\theta_0}^{\arccos\left(\frac{2}{3}\right)} \frac{d\theta}{\sin\left(\frac{\theta}{2}\right)}$ 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

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

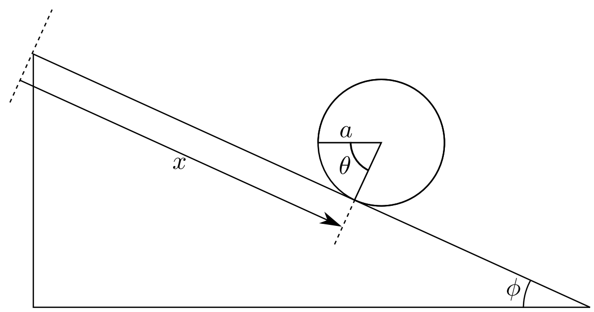


Figure 1.11: Hoop rolling on inclined plane

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

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

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

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

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

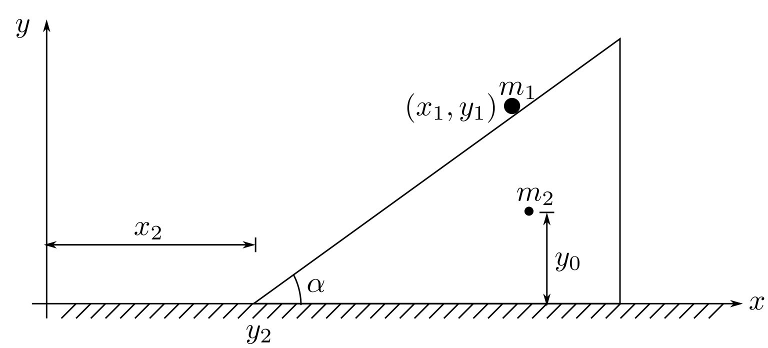


Figure 1.12: Wedge sliding on ice

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 $T = \frac{m_1}{2}\left(\dot{x}_1^2 + \dot{y}_1^2\right) + \frac{m_2}{2}\left(\dot{x}_2^2 + \dot{y}_2^2\right)$, 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

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

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

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

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

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

the accelerations are constant. As a check on our results, if $m_2 \to \infty$, 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 $\alpha \to 0$ 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:

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

Given these constraints, how many generalized coordinates are there?

If we know 3 non-collinear points in the body, the remaing 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 $N \to \infty$).

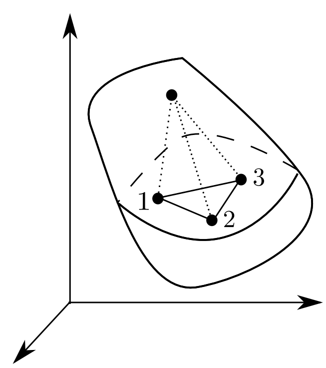


Figure 2.1: Three non-collinear points can be fully determined by using 6 continuous coordinates. Since the distances between any two other points are fixed in the rigid body, any other point of the body is fully determined by the distance to these 3 points (and a discrete orientation which is not a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

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

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:

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

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

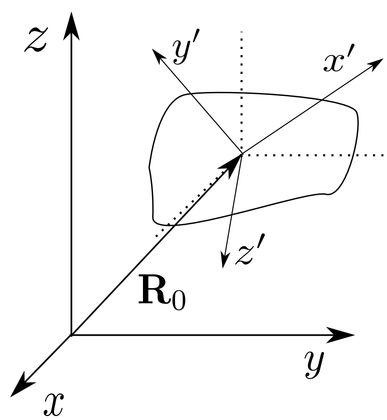


Figure 2.2: The three translational coordinates correspond to the position of the Center of Mass, and the three rotational coordinates correspond to the three angles necessary to define the orientation of the axis fixed with the body.

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.)

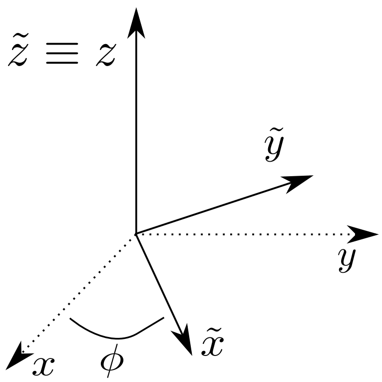


Figure 2.3: First rotation is by ϕ around the original z axis.

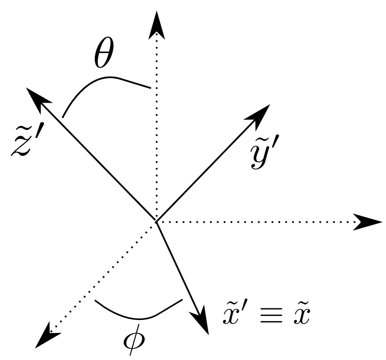


Figure 2.4: Second rotation is by θ around the intermediate $\tilde{x}$ axis.

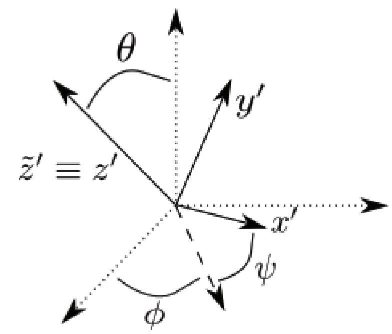


Figure 2.5: Final rotation is by ψ around the final z' axis.

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

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

Similar transformations can be written for the other terms:

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

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

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

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

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

All rotation matrices, including U_ϕ, U_θ, U_ψ, and U are *orthogonal*. Orthogonal matrices W satisfy

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

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:

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

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

$$U = \begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix} \qquad \text{or} \qquad U = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix} \tag{2.9}$$

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 $\det(W^\top W) = (\det(W))^2 = 1$. 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

$$U\mathbf{r} = \mathbf{r} \tag{2.10}$$

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

$$U\boldsymbol{\xi} = \lambda\boldsymbol{\xi}, \tag{2.11}$$

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

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

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

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

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

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

then we can rewrite Eq.(2.13) as

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

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:

$$U\boldsymbol{\xi} = \lambda\boldsymbol{\xi} \quad \Rightarrow \quad \boldsymbol{\xi}^\dagger U^\top = \lambda^\star \boldsymbol{\xi}^\dagger \quad \Rightarrow \quad \boldsymbol{\xi}^\dagger\boldsymbol{\xi} = \boldsymbol{\xi}^\dagger U^\top U \boldsymbol{\xi} = |\lambda|^2 \boldsymbol{\xi}^\dagger \boldsymbol{\xi}\,. \tag{2.16}$$

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

$$U\boldsymbol{\xi} = \lambda\boldsymbol{\xi} \Rightarrow U\boldsymbol{\xi}^\star = \lambda^\star\boldsymbol{\xi}^\star \tag{2.17}$$

where $\boldsymbol{\xi}^\star$ 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 $\lambda_2 = \lambda_3^\star$. 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 $\lambda_2 = \lambda_3^\star = e^{i\Phi}$ 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:

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

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 $\mathbf{e}'_i$.

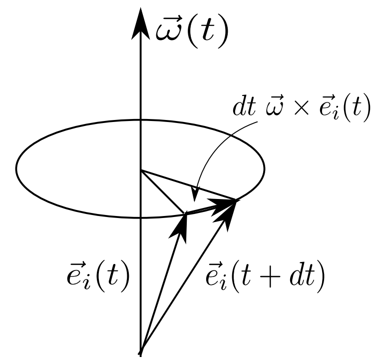
If we consider *any* vector then it can be decomposed with either set of basis vectors:

$$\mathbf{b} = \sum_i b_i \mathbf{e}_i = \sum_i b'_i \mathbf{e}'_i \,. \tag{2.19}$$

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

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

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 $\frac{d_{\mathrm{R}}}{dt}$ the time evolution in the rotating/body (R) frame, then vectors evolve in time according to

$$\frac{d\mathbf{b}}{dt} = \frac{d_{\mathrm{R}}\mathbf{b}}{dt} + \boldsymbol{\omega} \times \mathbf{b}. \tag{2.21}$$

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

$$\frac{d\mathbf{r}}{dt} = \frac{d_{\mathrm{R}}\mathbf{r}}{dt} + \boldsymbol{\omega} \times \mathbf{r} \quad \Leftrightarrow \quad \mathbf{v}_{\mathrm{F}} = \mathbf{v}_{\mathrm{R}} + \boldsymbol{\omega} \times \mathbf{r}. \tag{2.22}$$

Taking another time derivative gives us the analog for acceleration:

$$\begin{aligned} \frac{\mathbf{F}}{m} = \frac{d\mathbf{v}_{\mathrm{F}}}{dt} &= \frac{d_{\mathrm{R}}\mathbf{v}_{\mathrm{F}}}{dt} + \boldsymbol{\omega} \times \mathbf{v}_{\mathrm{F}} \\ &= \frac{d_{\mathrm{R}}\mathbf{v}_{\mathrm{R}}}{dt} + \frac{d_{\mathrm{R}}\boldsymbol{\omega}}{dt} \times \mathbf{r} + \boldsymbol{\omega} \times \frac{d_{\mathrm{R}}\mathbf{r}}{dt} + \boldsymbol{\omega} \times \mathbf{v}_{\mathrm{R}} + \boldsymbol{\omega} \times (\boldsymbol{\omega} \times \mathbf{r})\,. \end{aligned} \tag{2.23}$$

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

$$m\mathbf{a}_{\mathrm{R}} = \mathbf{F} - m\boldsymbol{\omega} \times (\boldsymbol{\omega} \times \mathbf{r}) - 2m\boldsymbol{\omega} \times \mathbf{v}_{\mathrm{R}} - m\dot{\boldsymbol{\omega}} \times \mathbf{r} \tag{2.24}$$

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 $-2m\boldsymbol{\omega} \times \mathbf{v}_{\mathrm{R}}$ are, respectively, the centrifugal and Coriolis ficticious forces respectively, while the last term $-m\frac{d_\mathrm{R}\boldsymbol{\omega}}{dt} \times \mathbf{r}$ is a ficticious 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:

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

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 $\boldsymbol{\omega} \times \boldsymbol{r} = \boldsymbol{\omega} \times \boldsymbol{r}_\perp$, so $-\boldsymbol{\omega} \times (\boldsymbol{\omega} \times \boldsymbol{r}_\perp) = w^2 \boldsymbol{r}_\perp$. 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: $\mathbf{v}_\mathrm{R} \neq 0$.

Example: Consider the impact of the Coriolis force on projectile motion on the rotating Earth, where the angular velocity is $\omega_\mathrm{Earth} = \frac{2\pi}{24 \times 3600\,\mathrm{s}} \approx 7.3 \times 10^{-5}\,\mathrm{s}^{-1}$. We work out the cross-product $-\boldsymbol{\omega} \times \boldsymbol{v}_r$ 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 $\boldsymbol{v}_r$.

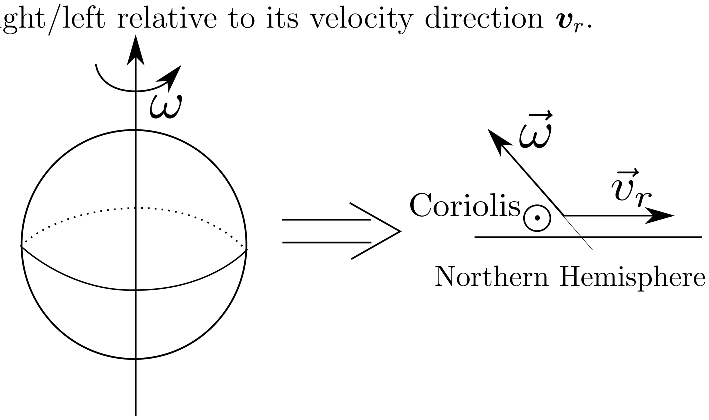
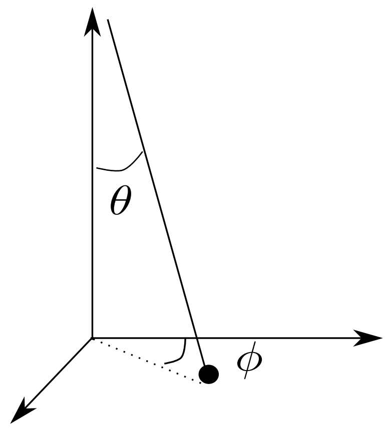
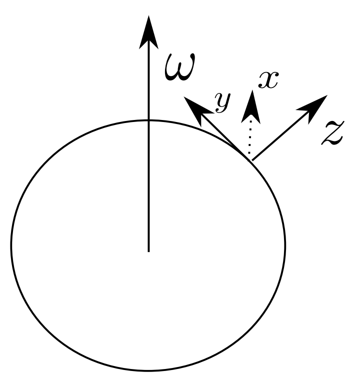


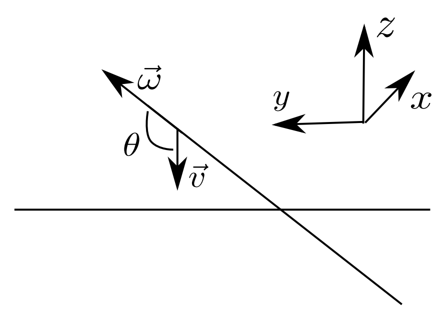
Figure 2.6: For a projectile, in the Northern Hemisphere, the Coriolis pushes it to its right, relative to its direction of motion.

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 $\boldsymbol{\omega}$, the result derived from the Coriolis force gives $\dot{\phi} \approx \omega_\mathrm{Earth} \sin(\lambda)$. 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 $\lambda = 90°$.



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 $\boldsymbol{\omega}$ gives us

$$m\mathbf{a}_{\mathrm{R}} = m\dot{\mathbf{v}}_{\mathrm{R}} = -mg\hat{z} - 2m\boldsymbol{\omega} \times \mathbf{v}, \tag{2.26}$$

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:

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

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

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

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

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

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

$$z = h_{\rm max} - \frac{g}{2}t^2 \tag{2.30}$$

and the mass reaches the floor at time

$$t_1 = \sqrt{\frac{2h_{\rm max}}{g}}. \tag{2.31}$$

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

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

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

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

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:

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

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:

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

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

$$T_{\rm R} = \frac{1}{2}m_i\dot{\mathbf{r}}_i^2 = \frac{1}{2}m_i(\boldsymbol{\omega}\times\mathbf{r}_i)^2 = \frac{1}{2}m_i\big[\boldsymbol{\omega}^2\mathbf{r}_i^2 - (\boldsymbol{\omega}\cdot\mathbf{r}_i)^2\big]. \tag{2.36}$$

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

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

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

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

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

$$\hat{I}_{ab} \equiv \int_{\mathcal{V}} (\mathbf{r}^2\delta_{ab} - r_a r_b)\rho(\mathbf{r})\, d\mathcal{V}. \tag{2.39}$$

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

$$T_{\mathrm{R}} = \frac{1}{2}\sum_{a,b} \hat{I}_{ab}\omega_a\omega_b = \frac{1}{2}\boldsymbol{\omega}\cdot\hat{I}\cdot\boldsymbol{\omega}\,, \tag{2.40}$$

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

$$\hat{I} = \sum_i m_i \begin{bmatrix} y_i^2+z_i^2 & -x_iy_i & -x_iz_i \\ -x_iy_i & x_i^2+z_i^2 & -y_iz_i \\ -x_iz_i & -y_iz_i & x_i^2+y_i^2 \end{bmatrix}, \tag{2.41}$$

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 $T_{\mathrm{R}} = \frac{1}{2}\hat{I}_{zz}\omega^2$ 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:

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

Writing out the components

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

which translates to the matrix equation:

$$\mathbf{L} = \hat{I} \cdot \boldsymbol{\omega}\,. \tag{2.44}$$

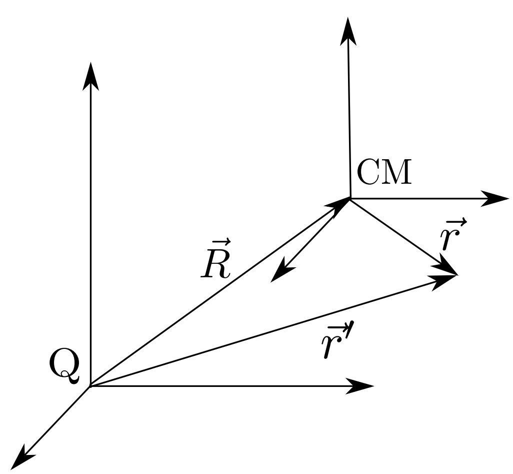
This allows us to write the corresponding rotational kinetic energy about the fixed point as:

$$T_{\rm R} = \frac{1}{2}\boldsymbol{\omega} \cdot \mathbf{L}. \tag{2.45}$$

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 $\hat{I}^{\rm CM}$ about the CM, it is simple to find $\hat{I}^{\rm Q}$ 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:

$$\mathbf{r}'_i = \mathbf{R} + \mathbf{r}_i\,. \tag{2.46}$$

By definition of the CM:

$$\sum_i m_i \mathbf{r}_i = 0 \qquad \text{and we let} \qquad M = \sum_i m_i. \tag{2.47}$$

The tensor of inertia around the new axis is then:

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

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

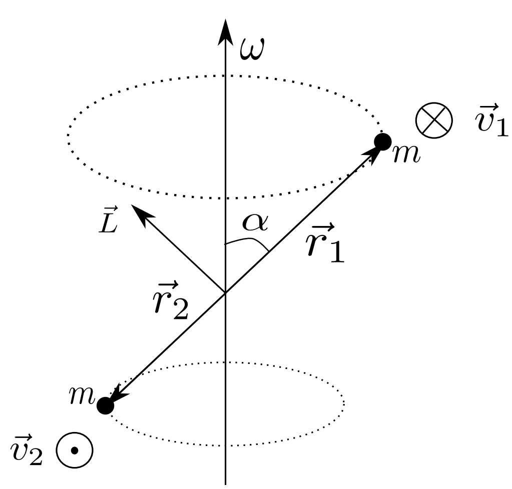
$$\hat{I}^{\mathrm{Q}}_{ab} = M(\delta_{ab}\mathbf{R}^2 - \boldsymbol{R}_a\boldsymbol{R}_b) + \hat{I}^{\mathrm{CM}}_{ab}, \tag{2.49}$$

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):

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

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 $\boldsymbol{L}$ is not parallel to $\boldsymbol{\omega}$. Consider a dumbbell made of 2 identical point passes 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,

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

Then using the definition of the moment inertia tensor:

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

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

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

Thus in this example $\boldsymbol{L}$ 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 $U^\top U = UU^\top = 1$. Vectors are rotated as

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

Putting these together

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

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 $\hat{I}'$:

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

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

$$\hat{I} \cdot \boldsymbol{\xi} = \lambda\, \boldsymbol{\xi}, \tag{2.57}$$

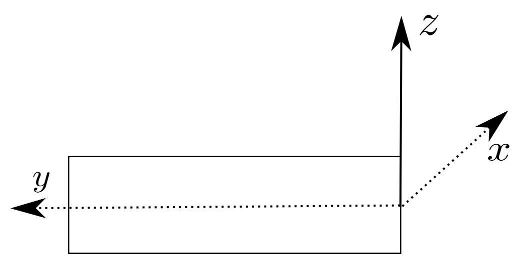
where the 3 eigenvalues λ give the principal moments of inertia I_α, and are obtained from solving $\det(\hat{I} - \lambda \mathbb{1}) = 0$. The corresponding 3 real *orthogonal* eigenvectors $\boldsymbol{\xi}^{(\alpha)}$ are the principal axes. Here $U^\top = \begin{bmatrix} \boldsymbol{\xi}^{(1)} &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

$$L_\alpha = I_\alpha \omega_\alpha \qquad \text{and} \qquad T = \frac{1}{2} \sum_\alpha I_\alpha\, \omega_\alpha^2\,, \tag{2.58}$$

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

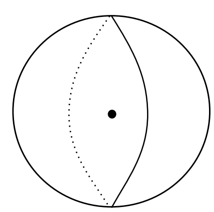
To summarize, for any choice of origin for any rigid body, there is a choice of axes that diagonalizes $\hat{I}$. For T to separate into translational and rotational parts, we must pick the origin to be the CM. Often, the principal axes can be identified by a symmetry of the body.

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.

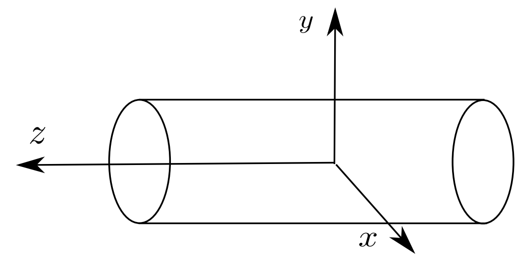


Sometimes, symmetry allows multiple choices for the principal axes.

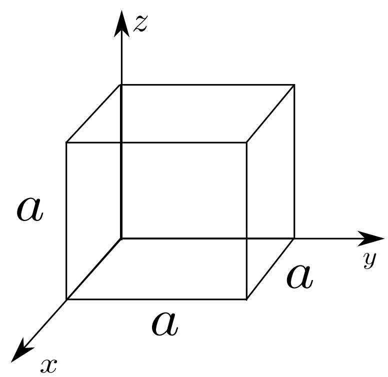
Example: for a sphere, any orthogonal axes through the origin are principal axes.



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

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

Thus the matrix is

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

The principal moments of inertia are found from

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

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

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

The remaining eigenvalues are degenerate:

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

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:

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

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

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

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:

$$\boldsymbol{\tau} = \frac{d\mathbf{L}}{dt} = \frac{d_{\text{R}}\mathbf{L}}{dt} + \boldsymbol{\omega} \times \mathbf{L} \tag{2.66}$$

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

$$\tau_1 = \frac{d_{\text{R}}L_1}{dt} + \omega_2 L_3 - \omega_3 L_2. \tag{2.67}$$

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*:

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

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

$$\tau_\alpha = I_\alpha\, \dot{\omega}_\alpha + \epsilon_{\alpha l k}\, \omega_l\, \omega_k\, I_k\,, \tag{2.69}$$

with α fixed but a sum implied over the repeated l and k indicies. 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

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

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 1[st] order. The Euler equations are:

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

so ω_1 is constant at this order, and

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

Combining these two equations yields

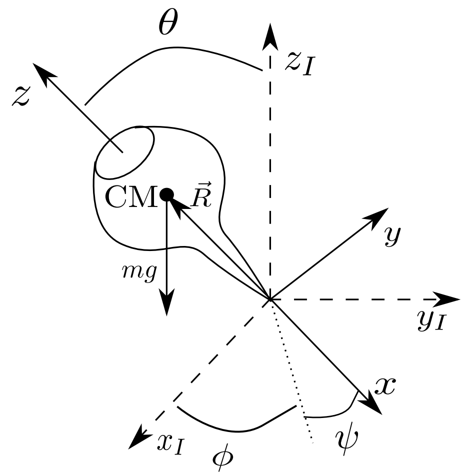
$$\ddot{\kappa}_2 = \left[\frac{(I_3 - I_1)(I_1 - I_2)\omega_1^2}{I_2 I_3}\right] \kappa_2\,, \tag{2.73}$$

and the same equation for κ_3. The term in the square bracket is constant, and is either negative $= -w^2$ with an oscillating solution $\kappa_2 \propto \cos(wt + \phi)$, or is positive $= \alpha^2$ with exponential solutions $\kappa_2 \propto ae^{\alpha t} + be^{-\alpha t}$. 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 $(x_\mathrm{I}, y_\mathrm{I}, z_\mathrm{I})$, 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



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

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

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

Let us define the constant:

$$
\Omega \equiv \frac{I_3 - I_1}{I_1}\omega_3\,. \tag{2.76}
$$

Then the Euler equations for this situation reduce to:

$$
\dot\omega_1 + \Omega\omega_2 = 0 \quad \text{and} \quad \dot\omega_2 - \Omega\omega_1 = 0\,. \tag{2.77}
$$

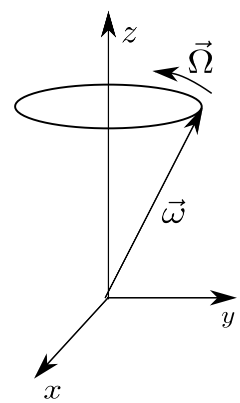
The simplest solution correspond to $\omega_1(t) = \omega_2(t) = 0$, where we just have a rotation about the z-axis. Here:

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

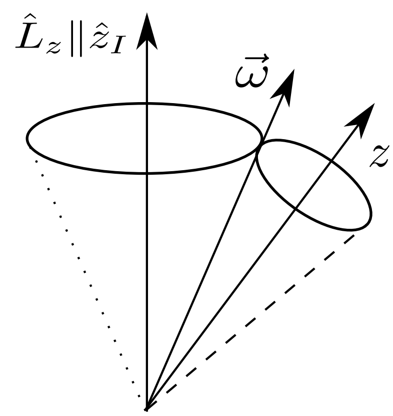
In this case $\mathbf{L} \parallel \boldsymbol{\omega}$. A more general situation is when $\mathbf{L}$ and $\boldsymbol{\omega}$ are not necessarily parallel, and ω_1 and ω_2 do not vanish. In this case Eq. (2.77) is solved by:

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

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 $T_{\mathrm{R}} = \frac{1}{2}\boldsymbol{\omega}\cdot\mathbf{L}$ 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 $\boldsymbol{L}$ 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:

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

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

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

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

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

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

$$\omega_3 = C\cot(\theta_0) - \Omega. \tag{2.83}$$

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:

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

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

Now we consider the general case where $\boldsymbol{\tau} \neq 0$ 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

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

gives us the kinetic energy

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

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

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

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

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 $\hat{z}_{\mathrm{I}}$ and $\hat{z}$). Additionally, the energy is given by

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

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:

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

Putting all this together gives:

$$\frac{\dot{u}^2}{2} = \left(\frac{2EI_3 - p_\psi^2}{2I_1I_3} - \frac{mgR}{I_1}u\right)\left(1-u^2\right) - \frac{1}{2}\left(\frac{p_\phi - p_\psi u}{I_1}\right)^2 \equiv -V_{\text{eff}}(u)\,, \tag{2.93}$$

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 $-1 \le (u = \cos\theta) \le 1$. 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.

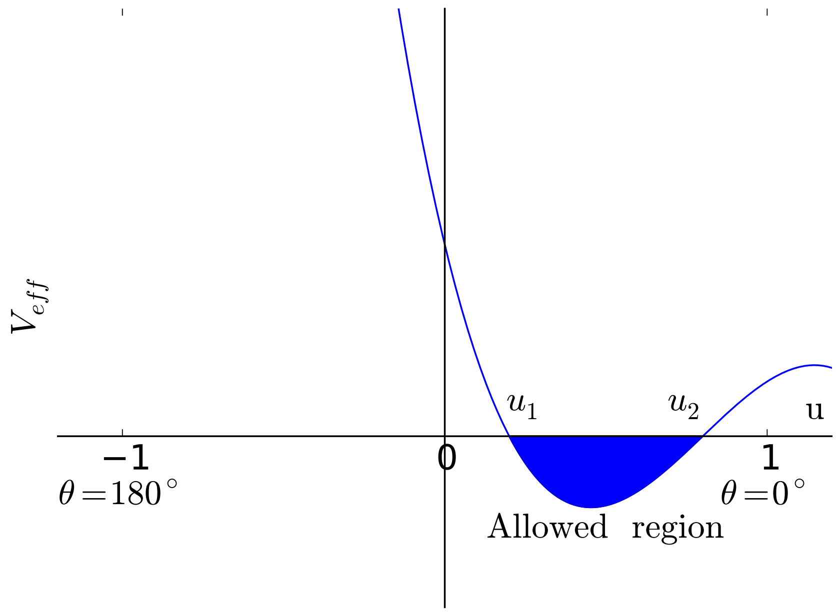


Figure 2.7: Allowed region for solutions for the top's nutation angle θ that solve Eq. (2.95).

Consider two roots u_1 and u_2 (corresponding respectively to some angles θ_1 and θ_2 as $u = \cos(\theta)$) satisfying $V_{\rm eff}(u_1) = V_{\rm eff}(u_2) = 0$, where $V_{\rm eff}(u) < 0$ 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, $\theta_1 = \theta_2$.

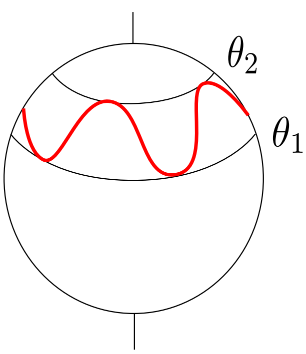


Figure 2.8: $\dot\phi > 0$

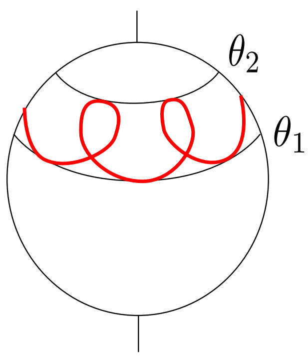


Figure 2.9: $\dot\phi$ has both signs

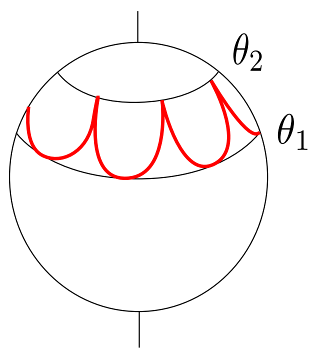


Figure 2.10: at θ_2 we have $\dot\phi = 0$, $\dot\theta = 0$

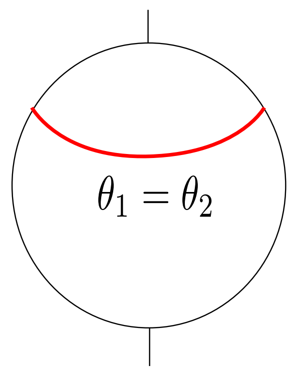


Figure 2.11: No nutation

Chapter 3

Vibrations & Oscillations

The topic of vibrations and oscillations is typically discussed in some detail in a course on waves (at MIT this is 8.03). Our goal for this chapter is to revisit aspects of oscillation phenomena using generalized coordinates. Many equations of motion we have encountered have been nonlinear. Here, we will expand about a minimum of the potential $V(q_1, \ldots, q_n)$, yielding linear equations.

Let us take $q_i = q_{0i} + \eta_i$, where $\vec{q}_0$ minimizes $V(q)$, and expand in the η_i. Henceforth and until further notice, repeated indices will implicitly be summed over. Then

$$V(q_1, \ldots, q_n) = V(q_{01}, \ldots, q_{0n}) + \left.\frac{\partial V}{\partial q_i}\right|_0 \eta_i + \frac{1}{2}\left.\frac{\partial^2 V}{\partial q_i \partial q_j}\right|_0 \eta_i \eta_j + \ldots, \tag{3.1}$$

where $|_0$ means "evaluate the quantity at $\vec{q}_0$". We already know that $\left.\frac{\partial V}{\partial q_i}\right|_0 = 0$ as by definition $\vec{q}_0$ minimizes $V(q)$. As a matter of convention, we choose $V(q_0) = 0$, since this just corresponds to picking the convention for the zero of the Energy. Finally, we define the constants $V_{ij} \equiv \left.\frac{\partial^2 V}{\partial q_i \partial q_j}\right|_0$. There is no time dependence in the definition of our generalized coordinates, so the kinetic energy is

$$T = \frac{1}{2} m_{ij}(q_1, \ldots, q_n) \dot{q}_i \dot{q}_j = \frac{1}{2} m_{ij}(q_{01}, \ldots, q_{0n}) \dot{\eta}_i \dot{\eta}_j + \mathcal{O}(\eta \dot{\eta}^2), \tag{3.2}$$

where $m_{ij}(q_{01}, \ldots, q_{0n}) \equiv T_{ij}$ are constants, and terms of $\mathcal{O}(\eta \dot{\eta}^2)$ and beyond are neglected. Thus, the Lagrangian to quadratic order in the η_is is

$$L = \frac{1}{2}\left(T_{ij} \dot{\eta}_i \dot{\eta}_j - V_{ij} \eta_i \eta_j\right) . \tag{3.3}$$

From this, the equations of motion are

$$T_{ij} \ddot{\eta}_j + V_{ij} \eta_j = 0 \tag{3.4}$$

to the same order. These are coupled linear equations of motion.

3.1 Simultaneous Diagonalization of $\hat{T}$ and $\hat{V}$

To solve Eq. (3.4) lets try

$$\eta_i = a_i e^{-\mathrm{i}\omega t} \tag{3.5}$$

where $a_i \in \mathbb{C}$ for all $i \in \{1, \ldots, n\}$, and the frequency ω is the same along all directions in the generalized coordinate space. Notationally, i and j will denote coordinate indices, while $\mathrm{i} = +\sqrt{-1}$ is the imaginary unit. This gives

$$V_{ij} a_j = \omega^2 T_{ij} a_j \tag{3.6}$$

which can be rewritten in matrix form as

$$\hat{V} \cdot \vec{a} = \lambda \hat{T} \cdot \vec{a} \tag{3.7}$$

with $\lambda = \omega^2$. This looks like an eigenvalue equation except that when we act with the linear operator $\hat{V}$ on $\vec{a}$ we get back $\hat{T} \cdot \vec{a}$ instead of just the eigenvector $\vec{a}$. This can be rewritten as

$$(\hat{V} - \lambda \hat{T}) \cdot \vec{a} = 0 \tag{3.8}$$

where $\hat{V}$ and $\hat{T}$ are *real* and *symmetric* $n \times n$ matrices. In order to have a non-trivial solution of this equation we need

$$\det(\hat{V} - \lambda \hat{T}) = 0 \tag{3.9}$$

which is an n^{th} order polynomial equation with n solutions eigenvalues λ_α with $\alpha \in \{1, \ldots, n\}$. The solutions of $(\hat{V} - \lambda_\alpha \hat{T}) \cdot \vec{a}^{(\alpha)} = 0$ are the eigenvectors $\vec{a}^{(\alpha)}$. This means

$$\hat{V} \cdot \vec{a}^{(\alpha)} = \lambda_\alpha \hat{T} \cdot \vec{a}^{(\alpha)}\,, \tag{3.10}$$

and the solutions are much like a standard eigenvalue problem. Here and henceforth, there will be no implicit sum over repeated eigenvalue indices α (so any sums that are needed will be made explicit), but we will retain implicit sums over repeated coordinate indices i & j.

There are two cases we will consider.

1) Let us start by considering the case when $\hat{T}$ is diagonal. In particular, let us consider the even easier case proportional to the unit matrix, where $T_{ij} = m\delta_{ij}$. This means

$$m\ddot{\eta}_i + V_{ij}\eta_j = 0\,. \tag{3.11}$$

Here we have the standard eigenvalue equation

$$\hat{V} \cdot \vec{a}^{(\alpha)} = m\lambda_\alpha \vec{a}^{(\alpha)}\,. \tag{3.12}$$

The eigenvalues λ_α are real and nonnegative as $\lambda_\alpha = \omega_\alpha^2$; the quantities ω_α are the *normal mode frequencies*. The eigenvectors $\vec{a}^{(\alpha)}$ are orthogonal, and we can choose their normalization so that

$$m\,\vec{a}^{(\beta)\dagger}\cdot\vec{a}^{(\alpha)} = \delta_{\beta\alpha} \qquad (\text{or } \vec{a}^{(\beta)\dagger}\cdot\vec{a}^{(\alpha)} = \delta_{\beta\alpha})\,. \tag{3.13}$$

This implies that

$$\lambda_\alpha = \vec{a}^{(\alpha)\dagger}\cdot\hat{V}\cdot\vec{a}^{(\alpha)} \qquad (\text{or } m\lambda_\alpha = \vec{a}^{(\alpha)\dagger}\cdot\hat{V}\cdot\vec{a}^{(\alpha)})\,. \tag{3.14}$$

The time-dependent eigenvectors are then

$$\vec{\eta}^{(\alpha)} = \vec{a}^{(\alpha)}e^{-\mathrm{i}\omega_\alpha t}, \quad \text{or} \quad \eta_i^{(\alpha)} = a_i^{(\alpha)}e^{-\mathrm{i}\omega_\alpha t}. \tag{3.15}$$

These are the normal mode solutions for the n coordinates labeled by i, and there are n such solutions labeled by α. The general solution of a linear equation is a superposition of the independent normal mode solutions:

$$\vec{\eta} = \sum_\alpha C_\alpha\,\vec{\eta}^{(\alpha)} \tag{3.16}$$

where $C_\alpha \in \mathbb{C}$ are fixed by initial conditions. To find real coordinate solutions, we take the real parts of these equations.

Lets prove the statements made above. Again, there will be no implicit sum over the eigenvalue index α. Dotting in $\vec{a}^{(\beta)\dagger}$ into Eq. (3.12) gives

$$\vec{a}^{(\beta)\dagger}\cdot\hat{V}\cdot\vec{a}^{(\alpha)} = m\lambda_\alpha\vec{a}^{(\beta)\dagger}\cdot\vec{a}^{(\alpha)}\,, \tag{3.17}$$

taking the Hermitian conjugate of both sides, noting that $\hat{V}^\dagger = \hat{V}$, and then swapping $\alpha \leftrightarrow \beta$ gives $\vec{a}^{(\beta)\dagger}\cdot\hat{V}\cdot\vec{a}^{(\alpha)} = m\lambda_\beta^\star\vec{a}^{(\beta)\dagger}\cdot\vec{a}^{(\alpha)}$. Taking the difference of these results gives

$$(\lambda_\alpha - \lambda_\beta^\star)\vec{a}^{(\beta)\dagger}\cdot\vec{a}^{(\alpha)} = 0, \tag{3.18}$$

and if $\alpha = \beta$ then $(\lambda_\alpha - \lambda_\alpha^\star)\vec{a}^{(\alpha)\dagger}\cdot\vec{a}^{(\alpha)} = 0$ implies the eigenvalues are real $\lambda_\alpha \in \mathbb{R}$. For $\lambda_\alpha \neq \lambda_\beta$, Eq. (3.18) then implies $\vec{a}^{(\beta)\dagger}\cdot\vec{a}^{(\alpha)} = 0$ so the eigenvectors are orthogonal. If by chance $\lambda_\alpha = \lambda_\beta$ for some $\alpha \neq \beta$ then we can always simply choose the corresponding eigenvectors to be orthogonal. By convention, we then normalize the eigenvectors so that they satisfy Eq. (3.13). Finally, if $\alpha = \beta$ then Eq. (3.17) now gives $\lambda_\alpha = \vec{a}^{(\alpha)\dagger}\cdot\hat{V}\cdot\vec{a}^{(\alpha)}$. The statement that we are at a local minimum of the multivariable potential and not a saddle point or a maximum implies then that $\lambda_\alpha \geq 0$ (we have positive second derivatives in each eigenvector direction).

2) Let us now consider when $\hat{T}$ is not diagonal and summarize which parts of the result are the same and where there are differences. Here we have $(\hat{V} - \lambda\hat{T})\cdot\vec{a} = 0$. Again, the eigenvalues λ_α are real and nonnegative, with $\lambda_\alpha = \omega_\alpha^2$. Now, however,

$$\vec{a}^{(\beta)\dagger}\cdot\hat{T}\cdot\vec{a}^{(\alpha)} = 0\,, \tag{3.19}$$

for $\alpha \neq \beta$, and we can replace the old normalization condition by a new one stating that

$$\vec{a}^{(\beta)\dagger} \cdot \hat{T} \cdot \vec{a}^{(\alpha)} = \delta_{\alpha\beta}, \tag{3.20}$$

which up to an overall prefactor reduces to the old orthonormality condition when $\hat{T} = m\hat{1}$. Here again,

$$\lambda_\alpha = \vec{a}^{(\alpha)\dagger} \cdot \hat{V} \cdot \vec{a}^{(\alpha)}\,, \tag{3.21}$$

and the α^{th} normal mode solution is

$$\vec{\eta}^{(\alpha)} = \vec{a}^{(\alpha)} e^{-\mathrm{i}\omega_\alpha t}\,. \tag{3.22}$$

The general solution is again the superposition

$$\vec{\eta} = \sum_\alpha C_\alpha \vec{\eta}^{(\alpha)}\,, \tag{3.23}$$

with the complex coefficients C_α fixed by the initial conditions (and a real part taken to get real coordinates).

Lets repeat the steps of our proof for this case. Dotting $\vec{a}^{(\beta)\dagger}$ into Eq. (3.10) gives

$$\vec{a}^{(\beta)\dagger} \cdot \hat{V} \cdot \vec{a}^{(\alpha)} = \lambda_\alpha \vec{a}^{(\beta)\dagger} \cdot \hat{T} \cdot \vec{a}^{(\alpha)}\,. \tag{3.24}$$

Taking the Hermitian conjugate and swapping α and β yields $\vec{a}^{(\beta)\dagger} \cdot \hat{V} \cdot \vec{a}^{(\alpha)} = \lambda_\beta^\star \vec{a}^{(\beta)\dagger} \cdot \hat{T} \cdot \vec{a}^{(\alpha)}$. Subtracting the two results this gives

$$(\lambda_\alpha - \lambda_\beta^\star)\, \vec{a}^{(\beta)\dagger} \cdot \hat{T} \cdot \vec{a}^{(\alpha)} = 0. \tag{3.25}$$

If we consider $\alpha = \beta$ then $(\lambda_\alpha - \lambda_\alpha^\star)\vec{a}^{(\alpha)\dagger} \cdot \hat{T} \cdot \vec{a}^{(\alpha)} = 0$ implies $\lambda_\alpha \in \mathbb{R}$ since $\hat{T} = \hat{T}^\star = \hat{T}^\top$ and physically we know that the kinetic energy $T = \dot{\vec{\eta}} \cdot \hat{T} \cdot \dot{\vec{\eta}} > 0$ for any $\dot{\vec{\eta}} \neq 0$. For $\lambda_\alpha \neq \lambda_\beta$, then the condition instead implies $\vec{a}^{(\beta)\dagger} \cdot \hat{T} \cdot \vec{a}^{(\alpha)} = 0$ so the eigenvectors are orthogonal (with respect to an inner product defined with $\hat{T}$). If by chance $\lambda_\alpha = \lambda_\beta$ for some $\alpha \neq \beta$ then we can simply choose the corresponding eigenvectors to be orthogonal in this sense. By convention, we normalize the eigenvectors so that they will be orthonormal as in Eq. (3.20). Finally, if $\alpha = \beta$ then $\lambda_\alpha = \vec{a}^{(\alpha)\dagger} \cdot \hat{V} \cdot \vec{a}^{(\alpha)}$, which is positive, so $\lambda_\alpha > 0$ also. The statement that we are at a local minimum of the potential and not a saddle point or a maximum implies then that $\lambda_\alpha \geq 0$.

3.2 Vibrations and Oscillations with Normal Coordinates

Given these results, it is natural to ask whether a different set of generalized coordinates might be better for studying motion about the minimum?

We form the matrix A by placing the eigenvectors in columns

$$A = \begin{bmatrix} \vec{a}^{(1)} & \vec{a}^{(2)} & \dots & \vec{a}^{(n)} \end{bmatrix} \tag{3.26}$$

and construct a diagonal eigenvalue matrix $\hat{\lambda} = \mathrm{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$. The matrix A can be ensured to be real because each $\vec{a}^{(\alpha)}$ only has at most an overall phase[1], and these can be removed by putting them into the coefficients C_α. The matrix A simultaneously diagonalizes $\hat{T}$ and $\hat{V}$ since

$$A^\top \hat{T} A = \mathbb{1} \quad \text{and} \quad A^\top \hat{V} A = \hat{\lambda}. \tag{3.27}$$

We choose new *normal coordinates* $\vec{\xi}$ by letting

$$\vec{\eta} = A\vec{\xi} \quad \text{and} \quad \dot{\vec{\eta}} = A\dot{\vec{\xi}} \tag{3.28}$$

so that the Lagrangian

$$\begin{aligned} L &= \frac{1}{2}\dot{\vec{\eta}} \cdot \hat{T} \cdot \dot{\vec{\eta}} - \frac{1}{2}\vec{\eta} \cdot \hat{V} \cdot \vec{\eta} \\ &= \frac{1}{2}\dot{\vec{\xi}} \cdot (A^\top \hat{T} A) \cdot \dot{\vec{\xi}} - \frac{1}{2}\vec{\xi} \cdot (A^\top \hat{V} A) \cdot \vec{\xi} \\ &= \frac{1}{2}\sum_\alpha \left(\dot{\xi}_\alpha^2 - \omega_\alpha^2 \xi_\alpha^2\right) . \end{aligned} \tag{3.29}$$

This gives the simple equations of motion for each α:

$$\ddot{\xi}_\alpha + \omega_\alpha^2 \, \xi_\alpha = 0 \, . \tag{3.30}$$

Thus, each normal coordinate describes the oscillations of the system with normal mode frequency ω_α.

Example: Let us consider the triatomic molecule CO_2 shown in Figure 3.1. We can picture it as a carbon atom of mass M in the middle of two oxygen atoms each of mass m. For the three particles there are 9 coordinates given by $\mathbf{r}_1$, $\mathbf{r}_2$, and $\mathbf{r}_3$. Six of these coordinates correspond to translations and rotations of the mass system treated as a rigid body. This leaves 3 coordinates that correspond to internal motions of the system. To model the potential we connect each oxygen atom to the carbon atom with a spring of constant k and relaxed length b. This does not add any cost to relative motion of the atoms with fixed spring length, which we will address below by adding another potential term in order to favor the linear configuration.

It is straightforward to guess what the normal modes could be:

[1] Why is it just an overall phase? The equation $\hat{V} \cdot \vec{a}^{(\alpha)} = \lambda_\alpha \hat{T} \cdot \vec{a}^{(\alpha)}$ alone does not fix the normalization of $\vec{a}^{(\alpha)}$. Let us say we pick $a_i^{(\alpha)} \in \mathbb{R}$ for some i. Then $V_{kj} a_j^{(\alpha)} = \lambda_\alpha K_{kj} a_j^{(\alpha)}$ is a set of equations with all real coefficients and one real term in the sums. Hence the solutions $a_j^{(\alpha)}/a_i^{(\alpha)} \in \mathbb{R}$ for all $j \in \{1, \dots, n\}$, implying that at most there is an overall phase in $\vec{a}^{(\alpha)}$.

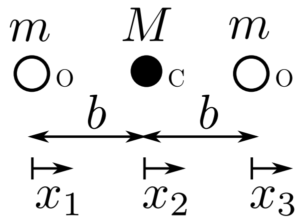


Figure 3.1: The CO_2 molecule.

- The oxygen atoms moving in the same direction along the line and the carbon atom moving in the opposite direction. This is a longitudinal oscillation.

- The oxygen atoms opposing each other along the line while the carbon atom remains at rest. This is a longitudinal oscillation.

- The oxygen atoms move in the same direction perpendicular to the line and the carbon atom moving in the opposite direction. This is a transverse oscillation.

These three normal modes are shown in Figure 3.2.

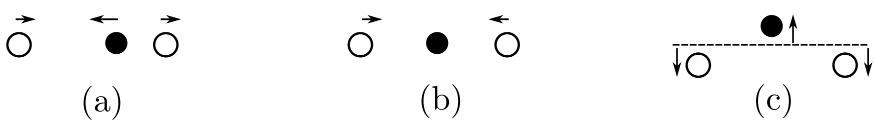


Figure 3.2: The three Normal Modes of the CO_2 molecule

We pick the body frame axes as follows:

- the three particles are in the xy-plane fixing 3 coordinates $z_i = 0$ for $i \in \{1, 2, 3\}$,

- the origin is the CM so $m(x_1 + x_3) + Mx_2 = m(y_1 + y_3) + My_2 = 0$, which fixes two more coordinates,

- the axes are oriented so that $y_1 = y_3$, which fixes one coordinate.

Defining the mass ratio as $\rho \equiv \frac{m}{M}$, then $x_2 = -\rho(x_1 + x_3)$ and $y_2 = -2\rho y_1$ can be eliminated. Altogether this fixes 6 coordinates, leaving the coordinates (x_1, x_3, y_1). This setup is shown in Figure 3.3.

For the potential we take

$$V = \frac{k}{2}\left(s_1 - b\right)^2 + \frac{k}{2}\left(s_2 - b\right)^2 + \frac{\lambda b^2}{2}(\alpha_1^2 + \alpha_2^2)\,. \tag{3.31}$$

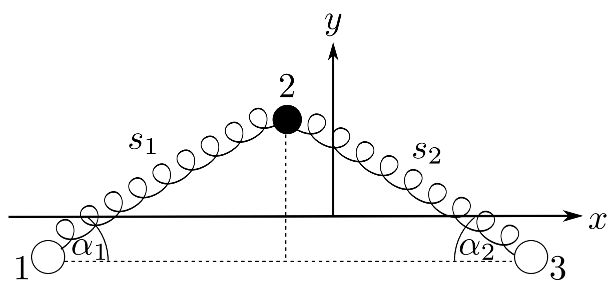


Figure 3.3: The orientation of the CO_2 molecule on xy plane

The first two terms are the springs discussed previously, and the last two provide a quadratic energy cost to the springs rotating away from the linear configuration, with strength given by λ. The spring lengths are

$$s_1 = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2} = \sqrt{[x_1 + \rho(x_1 + x_3)]^2 + (1 + 2\rho)^2 y_1^2} \tag{3.32}$$
$$s_2 = \sqrt{(x_2 - x_3)^2 + (y_2 - y_3)^2} = \sqrt{[x_3 + \rho(x_1 + x_3)]^2 + (1 + 2\rho)^2 y_1^2}\,,$$

and the two angles are

$$\alpha_1 = \tan^{-1}\left(\frac{y_3 - y_2}{x_3 - x_2}\right) = \tan^{-1}\left[\frac{(1 + 2\rho)y_1}{(1 + \rho)x_1 + x_3}\right], \tag{3.33}$$
$$\alpha_2 = \tan^{-1}\left(\frac{y_2 - y_1}{x_2 - x_1}\right) = \tan^{-1}\left[\frac{(1 + 2\rho)y_1}{(1 + \rho)x_3 + x_1}\right].$$

These results give $V = V(x_1, y_1, x_3)$.

For the kinetic energy we have

$$T = \frac{m}{2}\left(\dot{x}_1^2 + \dot{x}_3^2\right) + \frac{M}{2}\dot{x}_2^2 + \frac{m}{2}\left(\dot{y}_1^2 + \dot{y}_3^2\right) + \frac{M}{2}\dot{y}_2^2, \tag{3.34}$$

which after eliminating coordinates becomes

$$T = \frac{m}{2}(1 + \rho)\left(\dot{x}_1^2 + \dot{x}_3^2\right) + m\rho\dot{x}_1\dot{x}_3 + m(1 + 2\rho)\dot{y}_1^2. \tag{3.35}$$

Equilibrium comes from taking $y_1 = 0$, $x_3 = -x_1 = b$, which implies $\alpha_1 = \alpha_2 = 0$, $s_1 = s_2 = b$, and $V = 0$. We define coordinates for expanding about this potential minimum as $\eta_1 = x_1 + b$, $\eta_3 = x_3 - b$, and $\eta_2 = y_1$. Then as $V(-b, 0, b) = 0$ in equilibrium and $\left.\frac{\partial V}{\partial \eta_i}\right|_0 = 0$ then $V = \frac{1}{2}V_{ij}\eta_i\eta_j + \ldots$ where

$$V_{ij} = \left.\frac{\partial^2 V}{\partial \eta_i \partial \eta_j}\right|_0 = \begin{bmatrix} k(1 + 2\rho + 2\rho^2) & 0 & 2k\rho(1 + \rho) \\ 0 & 2\lambda(1 + 2\rho)^2 & 0 \\ 2k\rho(1 + \rho) & 0 & k(1 + 2\rho + 2\rho^2) \end{bmatrix} \tag{3.36}$$

for this system. Additionally,

$$T_{ij} = \begin{bmatrix} m(1+\rho) & 0 & m\rho \\ 0 & 2m(1+2\rho) & 0 \\ m\rho & 0 & m(1+\rho) \end{bmatrix} \tag{3.37}$$

for this system. Since there are no off-diagonal terms in the 2nd row or 2nd column in either $\hat{V}$ or $\hat{T}$, the transverse and the longitudinal modes decouple. For the transverse mode, we are left with

$$\ddot{y}_1 + \frac{2\lambda(1+2\rho)^2}{2m(1+2\rho)} y_1 = 0\,, \tag{3.38}$$

which is a simple harmonic oscillator. For the longitudinal modes, we have $\vec{\eta} = (\eta_1, \eta_3)$. The frequencies come from

$$\det \begin{bmatrix} k(1+2\rho+2\rho^2) - \lambda m(1+\rho) & 2k\rho(1+\rho) - \lambda m\rho \\ 2k\rho(1+\rho) - \lambda m\rho & k(1+2\rho+2\rho^2) - \lambda m(1+\rho) \end{bmatrix} = 0 \tag{3.39}$$

The solutions give the normal mode frequencies

$$\lambda_1 = \omega_1^2 = \frac{k}{m}\,, \qquad \lambda_2 = \omega_2^2 = \frac{k}{m}\left(1+2\rho\right)\,, \tag{3.40}$$

with associated eigenvectors

$$\vec{a}^{(1)} = \frac{1}{\sqrt{2m}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}\,, \qquad \vec{a}^{(2)} = \frac{1}{\sqrt{2m(1+2\rho)}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}\,, \tag{3.41}$$

which were chosen to satisfy $\vec{a}^{(\alpha)} \cdot \hat{T} \cdot \vec{a}^{(\beta)} = \delta_{\alpha\beta}$. Thus, the normal coordinates for the longitudinal modes are $\xi_1 \propto x_1 - x_3$ and $\xi_2 \propto x_1 + x_3$. Oscillations in these coordinates correspond to the normal mode motions in Fig. 3.2(b) and Fig. 3.2(a) respectively.

Chapter 4

Canonical Transformations, Hamilton-Jacobi Equations, and Action-Angle Variables

We've made good use of the Lagrangian formalism. Here we'll study dynamics with the Hamiltonian formalism. Problems can be greatly simplified by a good choice of generalized coordinates. How far can we push this?

Example: Let us imagine that we find coordinates q_i that are all cyclic. Then $\dot{p}_i = 0$, so $p_i = \alpha_i$ are all constant. If H is conserved, then:

$$H = H(\alpha_1, \ldots, \alpha_n) \tag{4.1}$$

is also constant in time. In such a case the remaining equations of motion:

$$\dot{q}_i = \frac{\partial H}{\partial \alpha_i} = \omega_i(\alpha) \quad \Rightarrow \quad q_i = \omega_i\, t + \delta_i \tag{4.2}$$

All coordinates are linear in time and the motion becomes very simple.

We might imagine searching for a variable transformation to make as many coordinates as possible cyclic. Before proceeding along this path, we must see what transformations are allowed.

4.1 Generating Functions for Canonical Transformations

Recall the the Euler-Lagrange equations are invariant when:

- a point transformation occurs $Q = Q(q,t)$ with $L[q,t] = L'[Q,t]$;
- a total derivative is summed to the Lagrangian $L' = L + \frac{dF[q,t]}{dt}$.

For H we consider point transformations in phase space:

$$Q_i = Q_i(q,p,t) \quad \text{and} \quad P_i = P_i(q,p,t), \tag{4.3}$$

where the Hamilton's equations for the evolution of the canonical variables (q,p) are satisfied:

$$\dot{q}_i = \frac{\partial H}{\partial p_i} \quad \text{and} \quad \dot{p}_i = -\frac{\partial H}{\partial q_i}. \tag{4.4}$$

Generally, not all transformations preserve the equations of motion. However, the transformation is *canonical* if there exists a new Hamiltonian:

$$K = K(Q,P,t), \tag{4.5}$$

where

$$\dot{Q}_i = \frac{\partial K}{\partial P_i} \quad \text{and} \quad \dot{P}_i = -\frac{\partial K}{\partial Q_i}. \tag{4.6}$$

For notational purposes let repeated indices be summed over implicitly.
Hamilton's principle can be written as:

$$\delta \int_{t_1}^{t_2} \left(p_i \dot{q}_i - H(q,p,t)\right) dt = 0, \tag{4.7}$$

or in the new Hamiltonian as:

$$\delta \int_{t_1}^{t_2} \left(P_i \dot{Q}_i - K(Q,P,t)\right) dt = 0. \tag{4.8}$$

For the Eq.(4.7) to imply Eq.(4.8), then we need:

$$\lambda(p_i \dot{q}_i - H) = P_i \dot{Q}_i - K + \dot{F}. \tag{4.9}$$

Since $\dot{F}$ is a total derivative and the ends of the path are fixed:

$$\delta q|_{t_1}^{t_2} = 0 \quad \text{and} \quad \delta p|_{t_1}^{t_2} = 0 \quad \Rightarrow \quad \delta F|_{t_1}^{t_2} = 0 \tag{4.10}$$

There are a few things to be said about transformations and λ.

- If $\lambda = 1$ then the transformation is *canonical*, which is what we will study.
- If $\lambda \neq 1$ then the transformation is *extended canonical*, and the results from $\lambda = 1$ can be recovered by rescaling q and p appropriately.

- If $Q_i = Q_i(q,p)$ and $P_i = P_i(q,p)$ without explicit dependence on time, then the transformation is *restricted canonical*.

We will always take transformations $Q_i = Q_i(q,p,t)$ and $P_i = P_i(q,p,t)$ to be invertible in any of the canonical variables. If F depends on a mix of old and new phase space variables, it is called a *generating function* of the canonical transformation. There are four important cases of this.

1. Let us take

$$F = F_1(q,Q,t) \tag{4.11}$$

where the old coordinates q_i and the new coordinates Q_i are independent. Then:

$$p_i\dot{q}_i - H = P_i\dot{Q}_i - K + \dot{F}_1 = P_i\dot{Q}_i - K + \frac{\partial F_1}{\partial t} + \frac{\partial F_1}{\partial q_i}\dot{q}_i + \frac{\partial F_1}{\partial Q_i}\dot{Q}_i \tag{4.12}$$

from this we see that $P_i\dot{Q}_i$ cancels and equating the terms with a $\dot{q}_i$, a $\dot{Q}_i$ and the remaining terms gives:

$$p_i = \frac{\partial F_1}{\partial q_i} \quad , \quad P_i = -\frac{\partial F_1}{\partial Q_i} \quad \text{and} \quad K = H + \frac{\partial F_1}{\partial t}, \tag{4.13}$$

which gives us formula for a transformation:

$$p_i = p_i(q,Q,t) \quad \text{and} \quad P_i = P_i(q,Q,t) \tag{4.14}$$

and connects K to an initial H.

Example: if

$$F_1 = -\frac{Q}{q}, \tag{4.15}$$

then:

$$p = \frac{\partial F_1}{\partial q} = \frac{Q}{q^2} \quad \text{and} \quad P = -\frac{\partial F_1}{\partial Q} = \frac{1}{q}. \tag{4.16}$$

Writing the new coordinates as function of the old ones yields

$$Q = pq^2 \quad \text{and} \quad P = \frac{1}{q} \tag{4.17}$$

Example: Given the transformations

$$Q = \ln\left(\frac{p}{q}\right) \quad \text{and} \quad P = -\left(\frac{q^2}{2} + 1\right)\frac{p}{q}, \tag{4.18}$$

we can prove they are canonical by finding a corresponding generating function. We know:

$$\frac{\partial F_1}{\partial q} = p = qe^Q, \tag{4.19}$$

which gives us

$$F_1 = \int qe^Q \, dq + g(Q) = \frac{q^2}{2}e^Q + g(Q), \tag{4.20}$$

and

$$\begin{aligned} P = -\frac{\partial F_1}{\partial Q} = -\frac{q^2}{2}e^Q - \frac{dg}{dQ} = -\left(\frac{q^2}{2}+1\right)\frac{p}{q} = -\left(\frac{q^2}{2}+1\right)e^Q \\ \Rightarrow g(Q) = e^Q\,. \end{aligned} \tag{4.21}$$

Thus F_1 is given by:

$$F_1 = \left(\frac{q^2}{2}+1\right)e^Q. \tag{4.22}$$

2. Let:

$$F = F_2(q,P,t) - Q_i P_i \tag{4.23}$$

where we wish to treat the old coordinates q_i and new momenta P_i as independent variables. Then:

$$\dot{q}_i p_i - H = \dot{Q}_i P_i - K + \dot{F}_2 - \dot{Q}_i P_i - Q_i \dot{P}_i = -Q_i \dot{P}_i - K + \frac{\partial F_2}{\partial t} + \frac{\partial F_2}{\partial q_i}\dot{q}_i + \frac{\partial F_2}{\partial P_i}\dot{P}_i \tag{4.24}$$

This corresponds to

$$p_i = \frac{\partial F_2}{\partial q_i} \quad ; \quad Q_i = \frac{\partial F_2}{\partial P_i} \quad \text{and} \quad K = H + \frac{\partial F_2}{\partial t}. \tag{4.25}$$

3. We could also take

$$F = F_3(p,Q,t) + q_i p_i \tag{4.26}$$

with the new coordinates Q_i and the old momenta p_i as independent variables.

4. Finally we could take

$$F = F_4(p,P,t) + q_i p_i - Q_i P_i \tag{4.27}$$

with the old momenta p_i and new momenta P_i as independent variables.

This can be summarized in the table below.

Function	Transformations	Simplest case	
$F_1(q,Q,t)$	$p_i = \frac{\partial F_1}{\partial q_i}, P_i = -\frac{\partial F_1}{\partial Q_i}$	$F_1 = q_i Q_i$	$Q_i = p_i, P_i = -q_i$
$F_2(q,P,t)$	$p_i = \frac{\partial F_2}{\partial q_i}, Q_i = \frac{\partial F_2}{\partial P_i}$	$F_2 = q_i P_i$	$Q_i = q_i, P_i = p_i$
$F_3(p,Q,t)$	$q_i = -\frac{\partial F_3}{\partial p_i}, P_i = -\frac{\partial F_3}{\partial Q_i}$	$F_3 = p_i Q_i$	$Q_i = -q_i, P_i = -p_i$
$F_4(p,P,t)$	$q_i = -\frac{\partial F_4}{\partial p_i}, Q_i = \frac{\partial F_4}{\partial P_i}$	$F_4 = p_i P_i$	$Q_i = p_i, P_i = -q_i$

The simplest case of the 2^{nd} (F_2) transformation is just an identity transformation. For any of these F_i cases we also have:

$$K = H + \frac{\partial F_i}{\partial t}. \tag{4.28}$$

If F_i is independent of time then this implies

$$K = H \tag{4.29}$$

Mixed cases may also occur when more than two old canonical coordinates are present. (In this chapter we will be using Einstein's repeated index notation for implicit summation, unless otherwise stated.)

Example: consider

$$F_2 = f_i(q,t)P_i \tag{4.30}$$

for some functions f_i where $i \in \{1,\ldots,n\}$. Then

$$Q_i = \frac{\partial F_2}{\partial P_i} = f_i(q,t) \tag{4.31}$$

is a coordinate point transformation. It is canonical with

$$p_i = \frac{\partial f_i}{\partial q_j} P_j, \tag{4.32}$$

which can be inverted to get $P_j = P_j(q,p,t)$.

Example: Consider the harmonic oscillator:

$$H = \frac{p^2}{2m} + \frac{kq^2}{2} \quad \text{where} \quad k = m\omega^2 \tag{4.33}$$

Let us try the transformation:

$$
\begin{aligned}
p &= \alpha\sqrt{2mP}\cos(Q) \\
q &= \frac{\alpha}{m\omega}\sqrt{2mP}\sin(Q)
\end{aligned}
\tag{4.34}
$$

for α constant. Then:

$$K = H = P\alpha^2\left(\cos^2(Q) + \sin^2(Q)\right) = P\alpha^2, \tag{4.35}$$

so the new momentum

$$P = \frac{E}{\alpha^2} \tag{4.36}$$

is just proportional to the energy, while Q is a cyclic variable.

Is this transformation canonical? We can find a generating function $F = F_1(q, Q)$ by dividing the old variables:

$$\frac{p}{q} = m\omega\cot(Q). \tag{4.37}$$

This gives us:

$$
\begin{aligned}
p &= \frac{\partial F_1}{\partial q} \Rightarrow F_1 = \int p(q,Q)\,dq + g(Q) = \frac{1}{2}m\omega q^2\cot(Q) + g(Q) \\
P &= -\frac{\partial F_1}{\partial Q} = \frac{m\omega q^2}{2\sin^2(Q)} - \frac{dg}{dQ}
\end{aligned}
\tag{4.38}
$$

Setting:

$$\frac{dg}{dQ} = 0 \Rightarrow q^2 = \frac{2P}{m\omega}\sin^2(Q), \tag{4.39}$$

which tells us the transformation is canonical if $\alpha = \sqrt{\omega}$. This means:

$$P = \frac{E}{\omega} \tag{4.40}$$

By Hamilton's equations Eq.(4.4):

$$\dot{Q} = \frac{\partial K}{\partial P} = \omega \Rightarrow Q = \omega t + \delta. \tag{4.41}$$

Putting this altogether, this gives the familiar results:

$$
\begin{aligned}
q &= \sqrt{\frac{2E}{m\omega^2}}\sin(\omega t + \delta) \\
p &= \sqrt{2mE}\cos(\omega t + \delta).
\end{aligned}
\tag{4.42}
$$

Lets record for future use our final canonical transformation here:

$$q = \sqrt{\frac{2P}{m\omega}}\sin(Q)\,, \qquad\qquad p = \sqrt{2m\omega P}\cos(Q)\,.$$

So far, a transformation $Q = Q(q, p, t)$ and $P = P(q, p, t)$ is canonical if we can find a generating function F. This involves integration, which could be complicated, so it would be nice to have a test that only involves differentiation. There is one!

4.2 Poisson Brackets and the Symplectic Condition

In Classical Mechanics II (8.223) the *Poisson bracket* of the quantities u and v was defined as

$$\{u, v\}_{q,p} = \sum_i \left(\frac{\partial u}{\partial q_i} \frac{\partial v}{\partial p_i} - \frac{\partial v}{\partial q_i} \frac{\partial u}{\partial p_i} \right) \tag{4.43}$$

It is easy to check that the following fundamental Poisson bracket relations are satisfied:

$$\{q_i, q_j\}_{q,p} = \{p_i, p_j\}_{q,p} = 0 \quad \text{and} \quad \{q_i, p_j\}_{q,p} = \delta_{ij}. \tag{4.44}$$

There are a few other properties of note. These include:

$$\{u, u\} = 0, \tag{4.45}$$

$$\{u, v\} = -\{v, u\}, \tag{4.46}$$

$$\{au + bv, w\} = a\{u, w\} + b\{v, w\}, \tag{4.47}$$

$$\{uv, w\} = u\{v, w\} + \{u, w\}v, \tag{4.48}$$

$$\{u, \{v, w\}\} + \{v, \{w, u\}\} + \{w, \{u, v\}\} = 0, \tag{4.49}$$

for a, b constants. Eq.(4.49) is the Jacobi identity.

The above looks a lot like the commutators of operators in quantum mechanics, such as:

$$[\hat{x}, \hat{p}] = \mathrm{i}\hbar \tag{4.50}$$

Indeed, quantizing a classical theory by replacing Poisson brackets with commutators through:

$$[u, v] = i\hbar\{u, v\} \tag{4.51}$$

is a popular approach (first studied by Dirac). It is also the root of the name "canonical quantization". (Note that Eq.(4.48) was written in a manner to match the analogous formula in quantum mechanics where the operator ordering is important, just in case its familiar. Here we can multiply functions in either order.)

Now we can state the desired criteria that only involves derivatives.

Theorem: A transformation $Q_j = Q_j(q, p, t)$ and $P_j = P_j(q, p, t)$ is canonical if and only if:

$$\{Q_i, Q_j\}_{q,p} = \{P_i, P_j\}_{q,p} = 0 \quad \text{and } \{Q_i, P_j\}_{q,p} = \delta_{ij}. \tag{4.52}$$

To prove it, we'll need some more notation. Let's get serious about treating q_i and p_i on an equal footing together, defining the following two quantities:

$$\vec{\eta} = \begin{pmatrix} q_1 \\ \vdots \\ q_n \\ p_1 \\ \vdots \\ p_n \end{pmatrix} \quad \text{and} \quad J = \begin{bmatrix} 0_{n\times n} & \mathbb{1}_{n\times n} \\ -\mathbb{1}_{n\times n} & 0_{n\times n} \end{bmatrix} \tag{4.53}$$

where $0_{n\times n}$ is the $n\times n$ zero matrix, $\mathbb{1}_{n\times n}$ is the $n\times n$ identity matrix. The following properties of J will be useful:

$$J^\top = -J \quad , \quad J^2 = -\mathbb{1}_{2n\times 2n} \quad \text{and} \quad J^\top J = JJ^\top = \mathbb{1}_{2n\times 2n}\,. \tag{4.54}$$

We also note that $\det(J) = 1$.

With this notation Hamilton's equations, Eq.(4.4), can be rewritten as:

$$\dot{\vec{\eta}} = J\frac{\partial H}{\partial \vec{\eta}} \quad \text{or} \quad \dot{\vec{\eta}} = J\nabla_{\vec{\eta}} H\,. \tag{4.55}$$

The notation $\nabla_{\vec{\eta}}H$ better emphasizes that this quantity is a vector, but we will stick to using the first notation for this vector, $\partial H/\partial\vec{\eta}$, below.

Although the Theorem is true for time dependent transformations, lets carry out the proof for the simpler case of time independent transformations $Q_i = Q_i(q,p)$ and $P_i = P_i(q,p)$. This implies $K = H$. Let us define:

$$\vec{\xi} = \begin{pmatrix} Q_1 \\ \vdots \\ Q_n \\ P_1 \\ \vdots \\ P_n \end{pmatrix} \tag{4.56}$$

which is a function of the original coordinates, so we can write:

$$\vec{\xi} = \vec{\xi}(\vec{\eta}) \tag{4.57}$$

Now consider the time derivative of $\vec{\xi}$:

$$\dot{\xi}_i = \frac{\partial \xi_i}{\partial \eta_j}\dot{\eta}_j \quad \Leftrightarrow \quad \dot{\vec{\xi}} = M\dot{\vec{\eta}} \quad \text{where} \quad M_{ij} = \frac{\partial \xi_i}{\partial \eta_j}. \tag{4.58}$$

Here M corresponds to the *Jacobian* of the transformation.

From the Hamilton's equations, we know that

$$\dot{\vec{\eta}} = J\frac{\partial H}{\partial \vec{\eta}}. \tag{4.59}$$

We want to show that :

$$\dot{\vec{\xi}} = J\frac{\partial H}{\partial \vec{\xi}} \quad \text{for} \quad \vec{\xi} = \vec{\xi}(\vec{\eta}) \text{ a canonical transformation.} \tag{4.60}$$

Let us now consider:

$$\dot{\xi}_i = \frac{\partial \xi_i}{\partial \eta_j}\dot{\eta}_j = \frac{\partial \xi_i}{\partial \eta_j} J_{jk} \frac{\partial H}{\partial \eta_k} = \frac{\partial \xi_i}{\partial \eta_j} J_{jk} \frac{\partial \xi_l}{\partial \eta_k}\frac{\partial H}{\partial \xi_l} \quad \Leftrightarrow \quad \dot{\vec{\xi}} = MJM^\top \frac{\partial H}{\partial \vec{\xi}} \tag{4.61}$$

for any H. Then

$$\vec{\xi} = \vec{\xi}(\vec{\eta}) \text{ is a } \textit{canonical transformation} \quad \text{iff} \quad MJM^\top = J \tag{4.62}$$

is satisfied. This is known as the "symplectic condition". Moreover, since

$$MJ = J(M^\top)^{-1} \quad \text{and} \quad J^2 = -\mathbb{1}, \tag{4.63}$$

we can write:

$$J(MJ)J = -JM = J(JM^{\top -1})J = -M^{\top -1}J \Rightarrow JM = M^{\top -1}J\,. \tag{4.64}$$

Thus we see the equivalence

$$MJM^\top = J \quad \Leftrightarrow \quad M^\top JM = J\,. \tag{4.65}$$

Now consider Poisson brackets in this matrix notation:

$$\{u, v\}_{q,p} = \{u, v\}_{\vec{\eta}} = \left(\frac{\partial u}{\partial \vec{\eta}}\right)^\top J\, \frac{\partial v}{\partial \vec{\eta}} \tag{4.66}$$

and the fundamental Poisson brackets are:

$$\{\eta_i, \eta_j\}_{\vec{\eta}} = J_{ij} \tag{4.67}$$

Then we can calculate the Poisson brackets that appeared in the theorem we are aiming to prove as

$$\{\xi_i, \xi_j\}_{\vec{\eta}} = \left(\frac{\partial \xi_i}{\partial \vec{\eta}}\right)^\top J \frac{\partial \xi_j}{\partial \vec{\eta}} = (M^\top JM)_{ij} \tag{4.68}$$

This last equation is the same as Eq.(4.65). The new variables satisfy the Poisson bracket relationships Eq.(4.67):

$$\{\xi_i, \xi_j\}_{\vec{\eta}} = J_{ij} \tag{4.69}$$

if and only if

$$M^\top JM = J \tag{4.70}$$

which using Eqs. (4.62) and (4.65) is true if and only if $\vec{\xi} = \vec{\xi}(\vec{\eta})$ is canonical. This completes the proof.

There are two facts that arise from this.

- Poisson brackets are canonical invariants

$$\{u, v\}_{\vec{\eta}} = \{u, v\}_{\vec{\xi}} = \{u, v\}. \tag{4.71}$$

This is true because:

$$\{u, v\}_{\vec{\eta}} = \left(\frac{\partial u}{\partial \vec{\eta}}\right)^{\top} J \frac{\partial v}{\partial \vec{\eta}} = \left(M^{\top} \frac{\partial u}{\partial \vec{\xi}}\right)^{\top} J \left(M^{\top} \frac{\partial v}{\partial \vec{\xi}}\right) \tag{4.72}$$

$$= \left(\frac{\partial u}{\partial \vec{\xi}}\right)^{\top} M J M^{\top} \frac{\partial v}{\partial \vec{\xi}} = \left(\frac{\partial u}{\partial \vec{\xi}}\right)^{\top} J \frac{\partial v}{\partial \vec{\xi}} = \{u, v\}_{\vec{\xi}} \tag{4.73}$$

- Phase space volume elements are preserved by canonical transformations, as discussed in 8.223. Phase space volume is given by:

$$V_{\vec{\xi}} = \prod_i dQ_i dP_i = |\det(M)| \prod_j dq_j dp_j = |\det(M)| V_{\vec{\eta}}. \tag{4.74}$$

Since we have

$$\det(M^{\top} J M) = \det(J) = (\det(M))^2 \det(J) \Rightarrow |\det(M)| = 1\,, \tag{4.75}$$

and the two phase space volumes are the same.

4.3 Equations of Motion & Conservation Theorems

Let us consider a function:

$$u = u(q, p, t) \tag{4.76}$$

Then:

$$\dot{u} = \frac{\partial u}{\partial q_i}\dot{q}_i + \frac{\partial u}{\partial p_i}\dot{p}_i + \frac{\partial u}{\partial t} = \frac{\partial u}{\partial q_i}\frac{\partial H}{\partial p_i} - \frac{\partial u}{\partial p_i}\frac{\partial H}{\partial q_i} + \frac{\partial u}{\partial t}, \tag{4.77}$$

which can be written more concisely as

$$\dot{u} = \{u, H\} + \frac{\partial u}{\partial t} \tag{4.78}$$

for any canonical variables (q, p) and corresponding Hamiltonian H. Performing canonical quantization on this yields the Heisenberg equation of time evolution in quantum mechanics. There are a few easy cases to check.

- If $u = q_i$ then:

$$\dot{q}_i = \{q_i, H\} + \frac{\partial q_i}{\partial t} = \frac{\partial H}{\partial p_i} \tag{4.79}$$

- If $u = p_i$ then:

$$\dot{p}_i = \{p_i, H\} + \frac{\partial p_i}{\partial t} = -\frac{\partial H}{\partial q_i} \tag{4.80}$$

  Together the above two cases yield Hamilton's equations of motion.

- Also, if $u = H$ then:

$$\dot{H} = \{H, H\} + \frac{\partial H}{\partial t} = \frac{\partial H}{\partial t} \tag{4.81}$$

Next consider what we can say if a quantity u is conserved. Here:

$$\dot{u} = 0 = \{u, H\} + \frac{\partial u}{\partial t}. \tag{4.82}$$

As a corollary, if

$$\frac{\partial u}{\partial t} = 0, \tag{4.83}$$

then

$$\{u, H\} = 0 \Rightarrow u \text{ is conserved.} \tag{4.84}$$

(In quantum mechanics this the analog of saying that u is conserved if u commutes with H.)

Another fact, is that if u and v are conserved then so is $\{u, v\}$. This could potentially provide a way to compute a new constant of motion. To prove it, first consider the special case where:

$$\frac{\partial u}{\partial t} = \frac{\partial v}{\partial t} = 0 \tag{4.85}$$

then using the Jacobi identity we have:

$$\{H, \{u, v\}\} = -\{u, \{v, H\}\} - \{v, \{H, u\}\} = -\{u, 0\} - \{v, 0\} = 0 \tag{4.86}$$

For the most general case we proceed in a similar manner:

$$\begin{aligned}\{\{u, v\}, H\} &= \{u, \{v, H\}\} + \{v, \{H, u\}\} = -\left\{u, \frac{\partial v}{\partial t}\right\} + \left\{v, \frac{\partial u}{\partial t}\right\} \\ &= -\frac{\partial}{\partial t}\{u, v\} \quad \Rightarrow \quad \frac{d}{dt}\{u, v\} = 0\end{aligned} \tag{4.87}$$

<u>Infinitesimal Canonical Transformations</u>

Let us now consider the generating function:

$$F_2(q, P, t) = q_i P_i + \epsilon G(q, P, t), \tag{4.88}$$

where $F_2 = q_i P_i$ is an identity transformation, and $|\epsilon| \ll 1$ is infinitesimal. The function $G(q, P, t)$ is known as the generating function of an infinitesimal canonical transformation. Using the properties of an F_2 generating function we have:

$$p_j = \frac{\partial F_2}{\partial q_j} = P_j + \epsilon \frac{\partial G}{\partial q_j} \quad \Rightarrow \quad \delta p_j = P_j - p_j = -\epsilon \frac{\partial G}{\partial q_j} \tag{4.89}$$

giving the infinitesimal transformation in the momentum. Likewise:

$$Q_j = \frac{\partial F_2}{\partial P_j} = q_j + \epsilon \frac{\partial G}{\partial P_j} \tag{4.90}$$

Since $P_j = p_j + O(\epsilon)$ and ϵ is infinitesimal we can replace $\partial G(q, P, t)/\partial P_j = \partial G(q, p, t)/\partial p_j + O(\epsilon)$. Therefore we have:

$$Q_j = q_j + \epsilon \frac{\partial G}{\partial p_j} + O(\epsilon^2) \quad \Rightarrow \quad \delta q_j = Q_j - q_j = \epsilon \frac{\partial G}{\partial p_j} \tag{4.91}$$

where now we note that we can consider $G = G(q, p, t)$, a function of q and p, to this order. Returning to the combined notation of $\vec{\eta}^\top = (q_1, \dots, q_n, p_1, \dots, p_n)$, Eq.(4.89) and Eq.(4.90) can be consisely written as the following Poisson bracket:

$$\delta \vec{\eta} = \epsilon \{\vec{\eta}, G\} \tag{4.92}$$

Example: if $G = p_i$ then $\delta p_i = 0$ and $\delta q_j = \epsilon \delta_{ij}$, which is why *momentum is the generator of spatial translations.*

Example: if G is the z component of the angular momentum:

$$G = L_z = \sum_i (x_i p_{iy} - y_i p_{ix}) \quad \text{and} \quad \epsilon = \delta\theta \tag{4.93}$$

then the infinitesimal change correponds to a rotation

$$\delta x_i = -y_i \delta\theta \quad , \quad \delta y_i = x_i \delta\theta \quad , \quad \delta z_i = 0 \tag{4.94}$$

$$\delta p_{ix} = -p_{iy} \delta\theta \quad , \quad \delta p_{iy} = p_{ix} \delta\theta \quad , \quad \delta p_{iz} = 0 \tag{4.95}$$

which is why *angular momentum is the generator of rotations.*

Important Example: if $G = H$ and $\epsilon = dt$ then

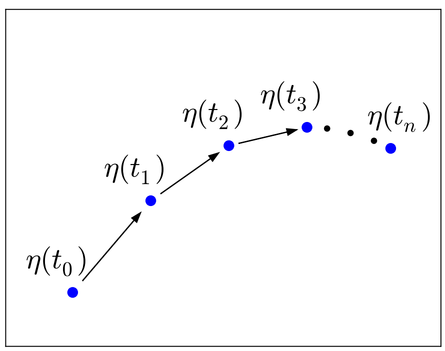
$$\epsilon \{\vec{\eta}, G\} = \{\vec{\eta}, H\} dt = \dot{\vec{\eta}} dt = d\vec{\eta}$$

On the left hand side we have the change to the phase space coordinates due to our transformation. On the right hand side we have the physical increment to the phase space variables that takes place in time dt. The fact that these are equivalent tells us that *the Hamiltonian is the generator of time evolution*. The infinitesimal transformation generated by the Hamiltonian gives the same function as solving for the physical motion.

Rather than trying to think of this as a passive coordinate change $(q, p) \to (Q, P)$, it is useful if we can take an *active view* of the infinitesimal canonical transformation generated by H. Let the time t be a parameter for the family of transformations with $\epsilon = dt$: the initial conditions are:

$$\vec{\eta}_0(t_0) = \vec{\eta}_0 \tag{4.96}$$

The result is a series of transformations of $\vec{\eta}$ that move us in a fixed set of phase space coordinates from one point to another:



Phase Space $\vec{\eta}$

$$\vec{\eta}_0(t_0) \to \vec{\eta}_1(\vec{\eta}_0, t_1) \to \ldots \to \vec{\eta}_n(\vec{\eta}_{n-1}, t_n)$$
$$\text{where } \ t_n = t \text{ is the final time} \tag{4.97}$$

All together, combining an infinite number of infinitesimal transformations allows us to make a finite transformation, resulting in:

$$\vec{\eta} = \vec{\eta}(\vec{\eta}_0, t) \qquad \text{or} \qquad \vec{\eta}_0 = \vec{\eta}_0(\vec{\eta}, t) \tag{4.98}$$

This is a canonical transformation that yields a solution for the motion!

How could we directly find this transformation, without resorting to stringing together infinitesimal transformations? We can simply look for a canonical transformation with new coordinates Q_i and new momenta P_i that are all constants, implying an equation of the type:

$$\vec{\eta}_0 = \vec{\eta}_0(\vec{\eta}, t) \tag{4.99}$$

Inverting this then gives the solution for the motion.

This logic can be used to extend our proof of the Theorem in Section 4.2 to fully account for time dependent transformations. (see Goldstein). Using $K = H + \epsilon \partial G/\partial t$, Goldstein also describes in some detail how the change to the Hamiltonian ΔH under an active infinitesimal canonical transformation satisfies:

$$\Delta H = -\epsilon\{G, H\} - \epsilon \frac{\partial G}{\partial t} = -\epsilon \dot{G} \tag{4.100}$$

This says "the constants of motion are generating functions G of the infinitesimal canonical transformation that leave H invariant"; that is, $\dot{G} = 0$ if and only if $\Delta H = 0$ under the transformation. Thus a conservation law exists if and only if there is a symmetry present.

4.4 Hamilton-Jacobi Equation

Let us take the suggestion from the end of the previous section seriously and look for new canonical variables that are all cyclic, such that:

$$\dot{Q}_i = \dot{P}_i = 0 \quad \Rightarrow \quad (Q, P) \text{ are all constants}\,. \tag{4.101}$$

If the new Hamiltonian K is independent of (Q, P) then:

$$\dot{Q}_i = \frac{\partial K}{\partial P_i} = 0 \quad \text{and} \quad \dot{P}_i = -\frac{\partial K}{\partial Q_i} = 0\,. \tag{4.102}$$

We could look for a constant K, but it it is simplest to simply look for $K = 0$.

Using a generating function $F = F_2(q, P, t)$, then we need

$$K = H(q, p, t) + \frac{\partial F_2}{\partial t} = 0\,. \tag{4.103}$$

Because $p_i = \dfrac{\partial F_2}{\partial q_i}$, then we can rewrite this condition as

$$H\left(q_1, \dots, q_n, \frac{\partial F_2}{\partial q_1}, \dots, \frac{\partial F_2}{\partial q_n}, t\right) + \frac{\partial F_2}{\partial t} = 0, \tag{4.104}$$

which is the time dependent *Hamilton-Jacobi equation* (henceforth abbreviated as the H-J equation). This is a 1^{st} order partial differential equation in $n+1$ variables $(q_1, \dots, q_n, t)$ for F_2. The solution for F_2 has $n+1$ independent constants of integration, One of these constants is trivial ($F_2 \to F_2 + C$ for a pure constant C), so we'll ignore this one. Hence, suppose the solution is:

$$F_2 \equiv S = S(q_1, \dots, q_n, \alpha_1, \dots \alpha_n, t), \tag{4.105}$$

where S is called Hamilton's principal function and each α_i is an independent constant. We can pick our new momenta to be the constants of integration $P_i = \alpha_i$ for $i \in \{1, \dots, n\}$ (so that $\dot{P}_i = 0$), thus specifying $F_2 = F_2(q, P, t)$ as desired. Then, using again the property of an F_2 generating function (and $K = 0$), we have that the new constant variables are:

$$P_i \equiv \alpha_i \quad \text{and} \quad Q_i \equiv \beta_i = \frac{\partial S(q, \alpha, t)}{\partial \alpha_i}. \tag{4.106}$$

We introduce the notation β_i to emphasize that these are constants.

From these results we can obtain a solution for the motion as follows. From the invertibility of our transformations we have:

$$\beta_i(q, \alpha, t) = \frac{\partial S}{\partial \alpha_i} \quad \Rightarrow \quad q_i = q_i(\alpha, \beta, t)\,, \tag{4.107}$$

$$p_i(q, \alpha, t) = \frac{\partial S}{\partial q_i} \quad \Rightarrow \quad p_i = p_i(q, \alpha, t) = p_i(q(\alpha, \beta, t), \alpha, t) = p_i(\alpha, \beta, t)\,.$$

(Note that function argument notation has been abused slightly here since $p_i(q, \alpha, t)$ and $p_i(\alpha, \beta, t)$ are technically different functions of their three arguments. Since we are always sticking explicit variables into the slots this should not cause confusion.) If desired, we can also swap our $2n$ constants α_i and β_i for $2n$ initial conditions q_{i0} and p_{i0}, to obtain a solution for the initial value problem. We obtain one set of constants in terms of the other set by solving the $2n$ equations obtained from the above results at $t = t_0$:

$$q_{i0} = q_i(\alpha, \beta, t_0)\,, \qquad\qquad p_{i0} = p_i(\alpha, \beta, t_0)\,. \tag{4.108}$$

Thus we see that Hamilton's principal function S is the generator of canonical transformations of constant (Q, P), and provides a method of obtaining solutions to classical mechanics problems by way of finding a transformation.

There are a few comments to be made about this.

1. The choice of constants α_i is somewhat arbitrary, as any other independent choice $\gamma_i = \gamma_i(\alpha)$ is equally good. Thus, when solving the H-J equation, we introduce the constants α_i in whatever way is most convenient.

2. What is S? We know that:

$$\dot{S} = \frac{\partial S}{\partial q_i}\dot{q}_i + \frac{\partial S}{\partial P_i}\dot{P}_i + \frac{\partial S}{\partial t}, \tag{4.109}$$

   but we also know that:

$$\frac{\partial S}{\partial q_i} = p_i \quad , \quad \dot{P}_i = 0 \quad \text{and} \quad \frac{\partial S}{\partial t} = -H \tag{4.110}$$

   Putting Eq.(4.109) and Eq.(4.110) together we have:

$$\dot{S} = p_i\dot{q}_i - H = L \quad \Rightarrow \quad S = \int L\,dt \tag{4.111}$$

   Thus S is the classical action which is an indefinite integral over time of the Lagrangian (so it is no coincidence that the same symbol is used).

3. The H-J equation is also the semiclassical limit of the quantum mechanical Schrödinger equation (0'th order term in the WKB approximation). To see this consider the Schrödinger equation

$$i\hbar\frac{\partial \psi}{\partial t} = H\psi = \Big(-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial q^2} + V(q)\Big)\psi\,, \tag{4.112}$$

with the wavefunction $\psi = \exp(iS/\hbar)$. At this point we are just making a change of variable, without loss of generality, and $S(q,t)$ is complex. Plugging it in, and canceling an exponential after taking the derivative, we find

$$-\frac{\partial S}{\partial t} = -\frac{i\hbar}{2m}\frac{\partial^2 S}{\partial q^2} + \frac{1}{2m}\Big(\frac{\partial S}{\partial q}\Big)^2 + V(q)\,. \tag{4.113}$$

This equation is just another way of writing the Schrödinger equation, to solve for a complex S instead of ψ. If we now take $\hbar \to 0$ then we find that the imaginary term goes away leaving

$$0 = \frac{\partial S}{\partial t} + \frac{1}{2m}\Big(\frac{\partial S}{\partial q}\Big)^2 + V(q) = \frac{\partial S}{\partial t} + H\Big(q, \frac{\partial S}{\partial q}\Big)\,, \tag{4.114}$$

which is the Hamilton-Jacobi equation for S with a standard $p^2/2m$ kinetic term in H.

Having set things up, it is always good for us to test a new formalism on an example where we know the solution.

Example: let us consider the harmonic oscillator Eq.(4.33):

$$H = \frac{1}{2m}\left(p^2 + (m\omega q)^2\right) = E \tag{4.115}$$

Here we will look for one constant $P = \alpha$ and one constant $Q = \beta$. The H-J equation says

$$\frac{1}{2m}\left(\left(\frac{\partial S}{\partial q}\right)^2 + (m\omega q)^2\right) + \frac{\partial S}{\partial t} = 0. \tag{4.116}$$

In solving this, we note that the dependence of S on q and t is *separable*

$$S(q,\alpha,t) = W(q,\alpha) + g(\alpha,t), \tag{4.117}$$

which gives:

$$\frac{1}{2m}\left(\left(\frac{\partial W}{\partial q}\right)^2 + (m\omega q)^2\right) = -\frac{\partial g}{\partial t} = \alpha\,. \tag{4.118}$$

Since the left side is independent of t and the right hand side is independent of q, then the result must be equal to a separation constant α that is independent of q and t. We will choose our new $P = \alpha$. Now we have

$$\frac{\partial g}{\partial t} = -\alpha \Rightarrow g = -\alpha t \tag{4.119}$$

where we have avoided the addition of a further additive constant (since our convention was to always drop an additive constant when determining S). To identify what α is note that

$$H = -\frac{\partial S}{\partial t} = -\frac{\partial g}{\partial t} = \alpha, \tag{4.120}$$

which corresponds to the constant energy,

$$\alpha = E. \tag{4.121}$$

The other equation we have to solve is

$$\frac{1}{2m}\left(\left(\frac{\partial W}{\partial q}\right)^2 + (m\omega q)^2\right) = \alpha, \tag{4.122}$$

so rearranging and integrating this gives the indefinite integral

$$W = \pm \int \sqrt{2m\alpha - (m\omega q)^2}\, dq\,, \tag{4.123}$$

which we will leave unintegrated until we must do so. The full solution is then given by:

$$S = -\alpha t \pm \int \sqrt{2m\alpha - (m\omega q)^2}\, dq. \tag{4.124}$$

With this result for Hamilton's Principal function in hand we can now solve for the equations of motion. The equations of motion come from (we now do the integral, after taking the partial derivative):

$$\beta = \frac{\partial S}{\partial \alpha} = -t \pm m \int \frac{dq}{\sqrt{2m\alpha - (m\omega q)^2}} \Rightarrow t + \beta = \pm\frac{1}{\omega}\arcsin\left(\sqrt{\frac{m\omega^2}{2\alpha}}q\right). \tag{4.125}$$

Inverting gives:

$$q = \pm\sqrt{\frac{2\alpha}{m\omega^2}}\sin(\omega(t+\beta)), \tag{4.126}$$

so β is related to the phase. Next we consider p and use this result to obtain:

$$p = \frac{\partial S}{\partial q} = \pm\sqrt{2m\alpha - (m\omega q)^2} = \pm\sqrt{2m\alpha}\cos(\omega(t+\beta)) \tag{4.127}$$

These results are as expected. We can trade (α, β) for the initial conditions (q_0, p_0) at $t = 0$. The choice of phase (from shifting β so that $\omega\beta \to \omega\beta + \pi$) allows taking the positive sign of each square root in the solutions above.

Separation of variables is the main technique to solve the H-J equation. In particular, for a time independent H where

$$\dot{H} = \frac{\partial H}{\partial t} = 0 \tag{4.128}$$

we can always separate time by taking:

$$S(q, \alpha, t) = W(q, \alpha) - \alpha_1\, t, \tag{4.129}$$

where α_1 has been chosen as the separation constant, then plugging this into the time dependent H-J equation yields (just as in our Harmonic Oscillator example):

$$H\left(q_1, \ldots, q_n, \frac{\partial W}{\partial q_1}, \ldots, \frac{\partial W}{\partial q_n}\right) = \alpha_1 \,. \tag{4.130}$$

This result is referred to as the *time independent Hamilton-Jacobi equation.* Since $\dot{H} = 0$, H is conserved, and equal to a constant α_1. If $H = E$ then this constant is energy, $\alpha_1 = E$. The function W is called Hamilton's characteristic function.

The idea is now to solve the time independent H-J equation for $W = W(q, \alpha)$ where $P = \alpha$ still. If we follow the setup from our time dependent solution above then the equations of motion are obtained from the following prescription for identifying variables:

$$\begin{aligned} p_i &= \frac{\partial W}{\partial q_i} \quad \text{for} \quad i \in \{1, \ldots, n\} \,, \\ Q_1 = \beta_1 &= \frac{\partial S}{\partial \alpha_1} = \frac{\partial W}{\partial \alpha_1} - t \,, \\ Q_j = \beta_j &= \frac{\partial W}{\partial \alpha_j} \quad \text{for } j \in \{2, \ldots, n\} \text{ for } n > 1 \,. \end{aligned} \tag{4.131}$$

Here all the Q_i are constants.

There is an alternative to the above setup, which allows us to not refer to the time dependent solution. The alternative is to consider $W = F_2(q, P)$ as the generating function, instead of S and only demand that all the new momenta P_i are constants with $P_1 = \alpha_1 = H$ for a time independent Hamiltonian H. At the start of chapter 4 we saw that this less restrictive scenario would lead to Qs that could have a linear time dependence, which is still pretty simple.

This is almost identical to the above setup but we rename and reshuffle a few things. The following three equations are the same as before:

$$p_i = \frac{\partial W}{\partial q_i} \quad , \quad P_i = \alpha_i \quad \text{and} \quad H\left(q, \frac{\partial W}{\partial q}\right) = \alpha_1 \tag{4.132}$$

However, now we have a non-zero K and different equation for Q_1:

$$K = H = \alpha_1 \quad \text{and} \quad Q_i = \frac{\partial W}{\partial \alpha_i} \text{ for all } i \in \{1, \ldots, n\}. \tag{4.133}$$

This means:

$$\dot{Q}_1 = \frac{\partial K}{\partial \alpha_1} = 1 \quad \Rightarrow \quad Q_1 = t + \beta_1 = \frac{\partial W}{\partial \alpha_1} \tag{4.134}$$

which is Eq. (4.131) but rearranged from the perspective of Q_1. For $j > 1$, the equations are the same as before Eq.(4.131):

$$\dot{Q}_j = \frac{\partial K}{\partial \alpha_j} = 0 \quad \Rightarrow \quad Q_j = \beta_j = \frac{\partial W}{\partial \alpha_j} \tag{4.135}$$

In this language we do not need to mention S only W. There are a few comments to be made:

1. Again, the choice of α is arbitrary, and $\alpha_i = \alpha_i(\gamma)$ is fine. If we do replace $\alpha_1 = \alpha_1(\gamma)$ then $\dot{Q}_i = \dfrac{\partial K}{\partial \gamma_i} = v_i$ is a constant so that (potentially) all of the Q_i become linear in time:
$$Q_i = v_i t + \beta_i \text{ for all } i \in \{1, \dots, n\} \tag{4.136}$$

2. What is W? We know that:
$$\dot{W} = \frac{\partial W}{\partial q_i}\dot{q}_i = p_i \dot{q}_i \Rightarrow W = \int p_i \dot{q}_i \, dt = \int p_i \, dq_i, \tag{4.137}$$
which is a different sort of "action".

3. The time independent H-J equation has some similarity to the time-independent Schrödinger energy eigenvalue equation (both involve H and and constant E, but the former is a non-linear equation for W, while the latter is a linear equation for the wavefunction ψ).

The H-J method is most useful when there is a separation of variables in H.

Example: if
$$H = h_1(q_1, q_2, p_1, p_2) + h_2(q_1, q_2, p_1, p_2) f(q_3, p_3) = \alpha_1, \tag{4.138}$$
so that q_3 is separable, then
$$f(q_3, p_3) = \frac{\alpha_1 - h_1}{h_2} \tag{4.139}$$
is a constant because the right hand side is independent of q_3 and p_3. Thus we assign
$$f(q_3, p_3) = \alpha_2 \tag{4.140}$$
for convenience. We can then write:
$$W = W'(q_1, q_2, \alpha) + W_3(q_3, \alpha) \quad \Rightarrow \quad f\left(q_3, \frac{\partial W_3}{\partial q_3}\right) = \alpha_2 \quad \text{and} \tag{4.141}$$
$$h_1\left(q_1, q_2, \frac{\partial W'}{\partial q_1}, \frac{\partial W'}{\partial q_2}\right) + \alpha_2 h_2\left(q_1, q_2, \frac{\partial W'}{\partial q_1}, \frac{\partial W'}{\partial q_2}\right) = \alpha_1 \tag{4.142}$$

Here, q_1 and q_2 may or may not be separable.

If all variables are separable then we use the solution:
$$W = \sum_i W_i(q_i, \alpha) \tag{4.143}$$

We can simply try a solution of this form to test for separability.

Note that cyclic coordinates are always separable.

Proof: let us say that q_1 is cyclic. Then

$$p_1 \equiv \gamma \quad \text{and} \quad H\left(q_2, \ldots, q_n, \frac{\partial W}{\partial q_1}, \frac{\partial W}{\partial q_2}, \ldots, \frac{\partial W}{\partial q_n}\right) = \alpha_1, \tag{4.144}$$

where γ is constant. Let us now write

$$W(q, \alpha) = W_1(q_1, \alpha) + W'(q_2, \ldots, q_n, \alpha). \tag{4.145}$$

This gives us:

$$p_1 = \frac{\partial W_1}{\partial q_1} = \gamma \Rightarrow W_1 = \gamma q_1. \tag{4.146}$$

Which gives us:

$$W(q, \alpha) = \gamma q_1 + W'(q_2, \ldots, q_n, \alpha) \tag{4.147}$$

This procedure can be repeated for any remaining cyclic variables.

Note that the choice of variables is often important in finding a result that separates. A problem with spherical symmetry may separate in spherical coordinates but not Cartesian coordinates.

4.5 Kepler Problem

As an extended example, let us consider the Kepler problem of two masses m_1 and m_2 in a central potential (with the CM coordinate $\mathbf{R} = 0$). The Lagrangian is:

$$L = \frac{1}{2} m \dot{\mathbf{r}}^2 - V(r) \quad \text{where} \quad \frac{1}{m} \equiv \frac{1}{m_1} + \frac{1}{m_2}, \tag{4.148}$$

and here m is the reduced mass. *Any* $V(r)$ conserves $\mathbf{L} = \mathbf{r} \times \mathbf{p}$, so the motion of $\mathbf{r}$ and $\mathbf{p}$ is in a plane perpendicular to $\mathbf{L}$. The coordinates in the plane can be taken as (r, ψ), so:

$$L = \frac{m}{2}\left(\dot{r}^2 + r^2 \dot{\psi}^2\right) - V(r), \tag{4.149}$$

with ψ being cyclic, which implies:

$$p_\psi = m r^2 \dot{\psi} \text{ is a constant.} \tag{4.150}$$

In fact $p_\psi = |\mathbf{L}| \equiv \ell$. Notationally, we use ℓ for the magnitude of the angular momentum $\mathbf{L}$ to distinguish it from the Lagrangian L.

The energy is then:

$$E = \frac{m}{2} \dot{r}^2 + \frac{\ell^2}{2 m r^2} + V(r), \tag{4.151}$$

which is constant, and this can be rewritten as:

$$E = \frac{m}{2}\dot{r}^2 + V_{\text{eff}}(r) \quad \text{where} \quad V_{\text{eff}}(r) = V(r) + \frac{\ell^2}{2mr^2}, \tag{4.152}$$

where V_{eff} is the effective potential, as plotted below for the gravitational potential.

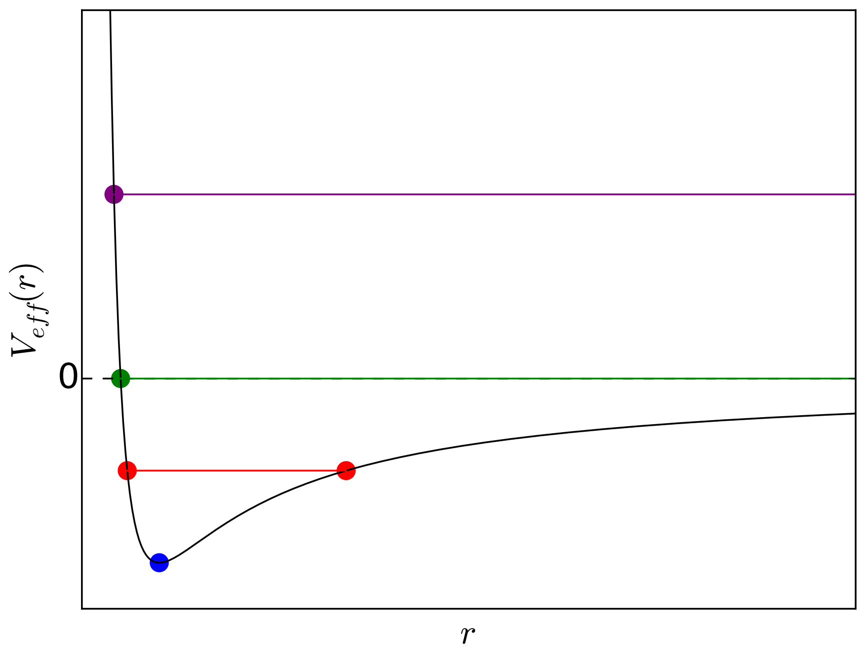


Figure 4.1: Plot of the effective potential V_{eff} along with the different qualitative orbits allowed in a gravity-like potential. The points correspond to turning points of the orbit.

Writing the E-L equation for $\dot{r} = dr/dt = \ldots$ and then solving for it as $dt = dr/(\ldots)$, and integrating yields

$$t = t(r) = \int_{r_0}^{r} \frac{dr'}{\sqrt{\frac{2}{m}\left(E - V(r') - \frac{\ell^2}{2mr'^2}\right)}} \tag{4.153}$$

as an implicit solution to the radial motion.

The orbital motion comes as $r = r(\psi)$ or $\psi = \psi(r)$ by using Eq.(4.150) and substituting, in Eq.(4.153). We have $\dot{\psi} = d\psi/dt = \ell/(mr^2)$, so we can use this to replace dt by $d\psi$ in $dt = dr/(\ldots)$ to get an equation of the form $d\psi = dr/(\ldots)$. The result is given by

$$\psi - \psi_0 = \ell \int_{r_0}^{r} \frac{dr'}{r'^2\sqrt{2m\left(E - V(r') - \frac{\ell^2}{2mr'^2}\right)}} \tag{4.154}$$

In the particular case of $V(r) = -\frac{k}{r}$, the solution of the orbital equation is:

$$\frac{1}{r(\psi)} = \frac{mk}{\ell^2}\left(1 + \varepsilon \cos(\psi - \psi')\right) \tag{4.155}$$

where the eccentricity ε is given by:

$$\varepsilon \equiv \sqrt{1 + \frac{2E\ell^2}{mk^2}} \tag{4.156}$$

Below are plotted the different qualitative orbits for the gravitic potential, with different ε or E (circular, elliptical, parabolic, and hyperbolic respectively).

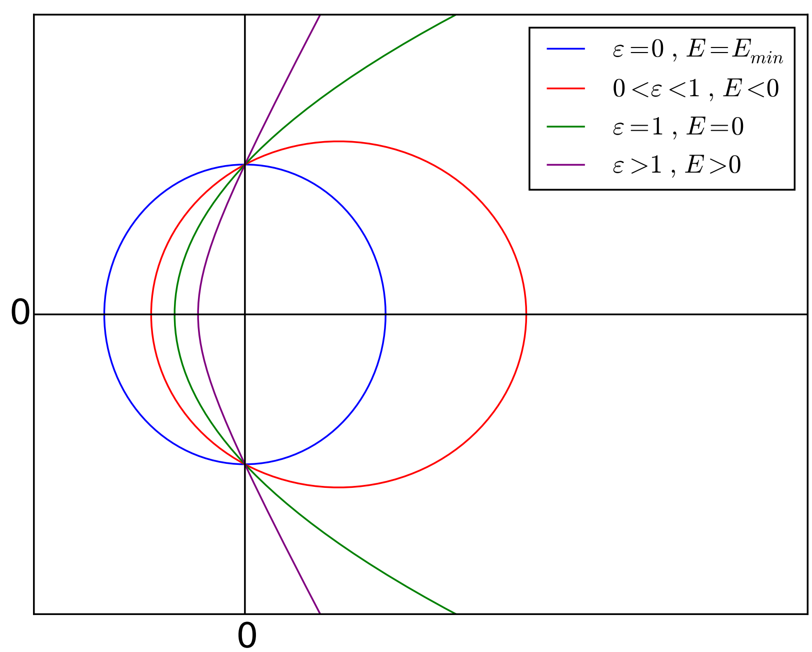


Figure 4.2: Different Orbits for the gravity-like potential. The orbits' colors match those of Fig.(4.1). The unbounded orbits occur for $E \geq 0$. The different curves correspond to the different possible conic sections.

Consider solving this problem instead by the H-J method. Lets start by considering as the variables (r, ψ) so that we assume that the motion of the orbit is in a plane. Here

$$H = \frac{1}{2m}\left(p_r^2 + \frac{p_\psi^2}{r^2}\right) + V(r) = \alpha_1 = E\,. \tag{4.157}$$

As ψ is cyclic, then $p_\psi \equiv \alpha_\psi$ is constant. Using:

$$W = W_1(r) + \alpha_\psi \psi, \tag{4.158}$$

then the time independent H-J equation is:

$$\frac{1}{2m}\left(\left(\frac{\partial W_1}{\partial r}\right)^2 + \frac{\alpha_\psi^2}{r^2}\right) + V(r) = \alpha_1. \tag{4.159}$$

This is simplified to

$$\frac{\partial W_1}{\partial r} = \sqrt{2m(\alpha_1 - V(r)) - \frac{\alpha_\psi^2}{r^2}} \tag{4.160}$$

and solved by

$$W = \alpha_\psi \psi + \int \sqrt{2m(\alpha_1 - V(r)) - \frac{\alpha_\psi^2}{r^2}}\, dr. \tag{4.161}$$

The transformation equations are:

$$\begin{aligned} t + \beta_1 &= \frac{\partial W}{\partial \alpha_1} = m \int \frac{dr}{\sqrt{2m(\alpha_1 - V(r)) - \frac{\alpha_\psi^2}{r^2}}}, \\ \beta_2 &= \frac{\partial W}{\partial \alpha_\psi} = \psi - \alpha_\psi \int \frac{dr}{r^2\sqrt{2m(\alpha_1 - V(r)) - \frac{\alpha_\psi^2}{r^2}}}. \end{aligned} \tag{4.162}$$

Thus we immediately get the radial equation $t = t(r)$ and orbital equation $\psi = \psi(r)$ from this, with $\alpha_\psi = \ell$ and $\alpha_1 = E$, showing that the constants are physically relevant parameters.

Let's solve this problem again, but suppose the motion is in 3 dimensions (as if we did not know the plane of the orbit). Using spherical coordinates (r, θ, φ) this corresponds to

$$H = \frac{1}{2m}\left(p_r^2 + \frac{p_\theta^2}{r^2} + \frac{p_\varphi^2}{r^2 \sin^2\theta}\right) + V(r) = \alpha_1\,. \tag{4.163}$$

Lets try a separable solution

$$W = W_r(r) + W_\theta(\theta) + W_\varphi(\varphi)\,. \tag{4.164}$$

Since φ is cyclic we know it is separable and that:

$$W_\varphi(\varphi) = \alpha_\varphi\, \varphi\,. \tag{4.165}$$

Together, this leaves;

$$\left(\frac{\partial W_r}{\partial r}\right)^2 + \frac{1}{r^2}\left(\left(\frac{\partial W_\theta}{\partial \theta}\right)^2 + \frac{\alpha_\varphi^2}{\sin^2(\theta)}\right) + 2mV(r) = 2m\alpha_1\,. \tag{4.166}$$

Because the term:

$$\left(\frac{\partial W_\theta}{\partial\theta}\right)^2 + \frac{\alpha_\varphi^2}{\sin^2(\theta)} \tag{4.167}$$

only depends on θ while the rest of the equation depends on r, it must be a constant so we can say:

$$\left(\frac{\partial W_\theta}{\partial\theta}\right)^2 + \frac{\alpha_\varphi^2}{\sin^2(\theta)} \equiv \alpha_\theta^2 \tag{4.168}$$

and the separation works. This then gives:

$$\left(\frac{\partial W_r}{\partial r}\right)^2 + \frac{\alpha_\theta^2}{r^2} = 2m(\alpha_1 - V(r)), \tag{4.169}$$

which is the same equation we considered before when assuming the motion was in a plane, with $\alpha_1 = E$ and $\alpha_\theta = \ell$. Eq. (4.168) says that

$$p_\theta^2 + \frac{p_\varphi^2}{\sin^2(\theta)} = \ell^2\,. \tag{4.170}$$

Here p_φ is the constant angular momentum about the $\hat{z}$ axis.

4.6 Action-Angle Variables

For many problems, we may not be able to solve analytically for the exact motion or for orbital equations, but we can still characterize the motion. For *periodic* systems we can find the frequency by exploiting *action-angle variables.*

The simplest case is for a single dimension of canonical coordinates (q, p). If $H(q, p) = \alpha_1$ then $p = p(q, \alpha_1)$. There are two types of periodic motion to consider.

1. Libration (oscillation) is characterized by a closed phase space orbit, so that q and p evolve periodically with the same frequency.

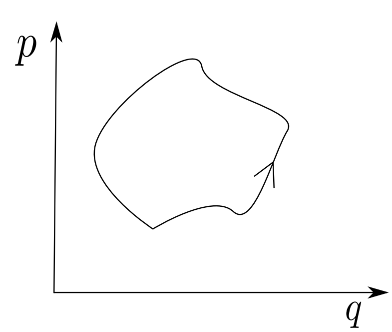


Figure 4.3: Phase space orbit of a libration (oscilation). The trajectory closes on itself, the state returns to the same position after some time τ.

2. Rotation is characterized by an open phase space path, so p is periodic while q evolves without bound.

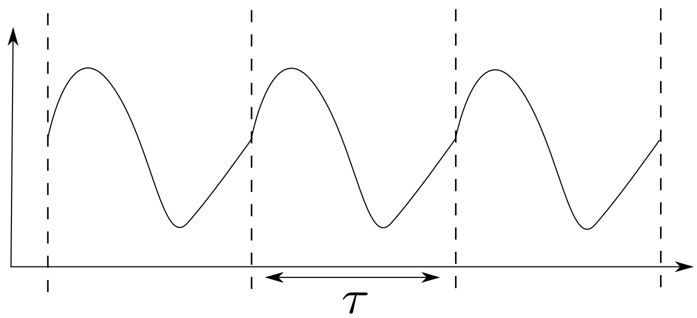


Figure 4.4: Phase space orbit of a rotation. Although the orbit is not closed, each period the evolution of the system is the same, leading to a orbit that repeats itself with a translation.

Example: a pendulum of length a may be characterized by canonical coordinates (θ, p_θ), where:

$$E = H = \frac{p_\theta^2}{2ma^2} - mga\cos\theta \tag{4.171}$$

This means:

$$p_\theta = \pm\sqrt{2ma^2(E + mga\cos\theta)} \tag{4.172}$$

must be real. A rotation occurs when $E > mga$, and oscillations occur when $E < mga$. The critical point in between (when the pendulum just makes it to the top) is when $E = mga$ exactly, and is depicted by a dashed line in the figure below.

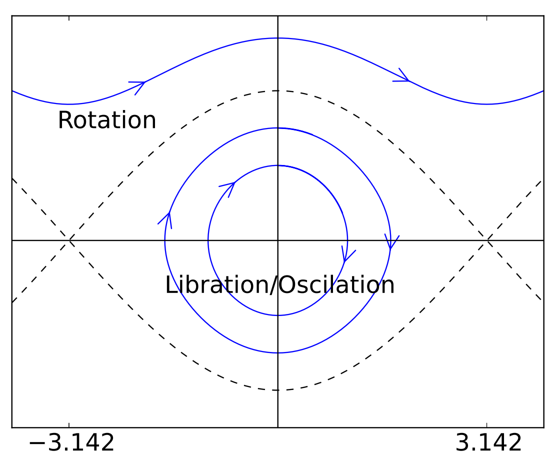


Figure 4.5: The pendulum exhibits both librations and rotations depending on the initial conditions.

For either type of periodic motion, it is useful to replace $P = \alpha_1$ by a different (constant) choice called the *action variable*

$$J = \oint p\, dq, \tag{4.173}$$

where $\oint$ refers to a definite integral over one period in phase space. To see that J is constant, recall that $p = p(q, \alpha_1)$, so plugging this into the definite integral we are left with $J = J(\alpha_1)$. Also, we have the inverse, $\alpha_1 = H = H(J)$, and can rewrite Hamilton's characteristic function in terms of J by $W = W(q, \alpha_1) = W(q, H(J)) = W(q, J)$ (where again the argument notation is abused slightly).

The coordinate conjugate to J is the *angle variable*

$$\omega = \frac{\partial W}{\partial J} \tag{4.174}$$

(where ω is not meant to imply an angular velocity). This means

$$\dot{\omega} = \frac{\partial H(J)}{\partial J} = \nu(J) \text{ is a constant.} \tag{4.175}$$

As a result the angle variable has linear time dependence,

$$\omega = \nu t + \beta\,, \tag{4.176}$$

for some initial condition β. Dimensionally, J has units of angular momentum, while ω has no dimensions (like an angle or a phase).

To see why it is useful to use the canonical variables (ω, J), let us consider the change in ω when q goes through a complete cycle.

$$\Delta\omega = \oint \frac{\partial \omega}{\partial q}\, dq = \oint \frac{\partial^2 W}{\partial q \partial J}\, dq = \frac{\partial}{\partial J} \oint \frac{\partial W}{\partial q}\, dq = \frac{\partial}{\partial J} \oint p\, dq = 1 \tag{4.177}$$

where in the last equality we used the definition of J in Eq.(4.173). Also, we have $\Delta\omega = \nu\tau$ where τ is the period. Thus

$$\nu = \frac{1}{\tau} \tag{4.178}$$

is the *frequency* of periodic motion. If we find $H = H(J)$ then

$$\nu = \frac{\partial H(J)}{\partial J} \tag{4.179}$$

immediately gives the frequency $\nu = \nu(J)$ for the system. Often, we then $J = J(E)$ to get $\nu = \nu(E)$ the frequency at a given energy. This is a very efficient way of finding the frequency of the motion without solving for extraneous information.

Example: let us consider a pendulum with action-angle variables. We define:

$$\tilde{E} = \frac{E}{mga} \tag{4.180}$$

so that $\tilde{E} > 1$ corresponds to rotation and $\tilde{E} < 1$ corresponds to oscillation. This means

$$p_\theta = \pm\sqrt{2m^2ga^3}\sqrt{\tilde{E} + \cos\theta}. \tag{4.181}$$

For $\tilde{E} > 1$:

$$J = \sqrt{2m^2ga^3}\int_{-\pi}^{\pi} d\theta\, \sqrt{\tilde{E} + \cos\theta}\,, \tag{4.182}$$

which is an elliptic integral. For $\tilde{E} < 1$:

$$\begin{aligned} J &= \sqrt{2m^2ga^3}\int_{-\theta_0}^{+\theta_0} d\theta\, \sqrt{\tilde{E} + \cos\theta} + \sqrt{2m^2ga^3}\int_{\theta_0}^{-\theta_0} d\theta \left[-\sqrt{\tilde{E} + \cos\theta}\,\right] \\ &= 4\sqrt{2m^2ga^3}\int_0^{\theta_0} d\theta\, \sqrt{\tilde{E} + \cos\theta}\,, \end{aligned} \tag{4.183}$$

as the contributions from the four intervals that the pendulum swings through in one period are all equivalent. Here θ_0 is the turning point of the oscillation, and $\tilde{E} = -\cos(\theta_0)$.

From this:

$$\nu = \frac{\partial E}{\partial J} = \left(\frac{\partial J}{\partial E}\right)^{-1} \tag{4.184}$$

which we can solve graphically by making a plot of J vs E, then dJ/dE versus E, and finally the inverse $\nu = dE/dJ$ versus E.

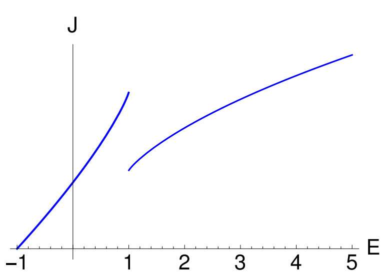


Figure 4.6: Plot of $J(E)$ versus $\tilde{E}$. The discontinuity corresponds to the transition from Oscilation to Rotation.

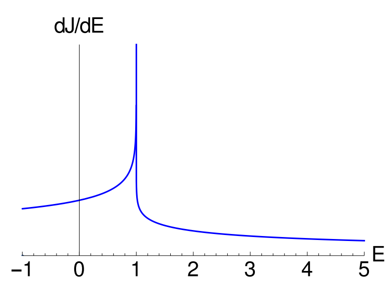


Figure 4.7: Plot of $\dfrac{dJ}{dE}$. The discontinuity is logarithmic divergent so it is integrable.

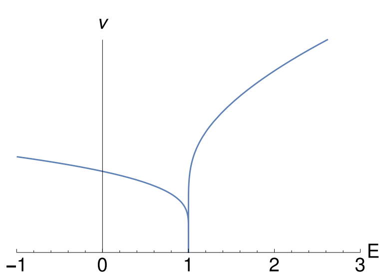


Figure 4.8: Plot of the frequency of oscillation $\nu(E)$ versus $\tilde{E}$. As $\tilde{E} \to -1$ we approach the small amplitude limit, where $\nu = (2\pi)^{-1}\sqrt{g/a}$.

Example: let us consider the limit $|\theta| \ll 1$ of a pendulum, so:

$$H = \frac{p_\theta^2}{2ma^2} + \frac{mga}{2}\theta^2 - mga \tag{4.185}$$

We can actually consider this in the context of a general harmonic oscillator

$$H = \frac{p^2}{2m} + \frac{m\Omega^2}{2} x^2, \tag{4.186}$$

where:

$$\Omega = \sqrt{\frac{g}{a}} \quad , \quad x = a\theta \quad \text{and} \quad p = \frac{p_\theta}{a} \tag{4.187}$$

Notationally, Ω is used for the harmonic oscillator frequency to distinguish from the transformed angle variable ω. We then have:

$$J = \oint p\, dq = \oint \pm\sqrt{2mE - m^2\Omega^2 x^2}\, dx \tag{4.188}$$

Note that the coordinate does not need to be an angle, as may be the case for general x. This gives:

$$J = 4\sqrt{2mE} \int_0^{x_0} \sqrt{1 - \frac{x^2}{x_0^2}}\, dx \quad \text{where} \quad x_0 \equiv \frac{\sqrt{2mE}}{m\Omega}. \tag{4.189}$$

Solving the integral yields

$$J = \pi\sqrt{2mE}x_0 = \frac{2\pi mE}{m\Omega} = \frac{2\pi E}{\Omega}, \tag{4.190}$$

which gives us

$$\frac{\partial E}{\partial J} = \frac{\Omega}{2\pi}, \tag{4.191}$$

the expected cyclic frequency for the harmonic oscillator.

Multiple Variables: We can treat multiple variables $(q_1, \ldots, q_n, p_1, \ldots, p_n)$ with the action-angle formalism if *each* pair (q_i, p_i) has an oscillatory or rotating orbit. Lets also assume that the H-J equation is completely separable into:

$$W = \sum_j W_j(q_j, \alpha). \tag{4.192}$$

Here we have

$$p_i = \frac{\partial W_i}{\partial q_i} = p_i(q_i, \alpha_1, \ldots, \alpha_n) \Rightarrow J_i = \oint p_i\, dq_i = J_i(\alpha_1, \ldots, \alpha_n) \tag{4.193}$$

where repeated indices do not correspond to implicit sums here. This implies that the inverse will be $\alpha_j = \alpha_j(J_1, \ldots, J_n)$ and thus $\alpha_1 = H = H(J_1, \ldots, J_n)$. Likewise:

$$\omega_i = \frac{\partial W}{\partial J_i} = \sum_j \frac{\partial W_j}{\partial J_i} = \omega_i(q_1, \ldots, q_n, J_1, \ldots, J_n)\,. \tag{4.194}$$

Just as in the one dimensional case the time derivative of the angle variables is a constant

$$\dot{\omega}_i = \frac{\partial H}{\partial J_i} = \nu_i(J_1, \ldots, J_n) \tag{4.195}$$

which are the frequencies describing motion in this "multi-periodic" system. Due to the presence of multiple frequencies, the motion through the whole $2n$-dimensional phase space need not be periodic in time.
Example: in the 2-dimensional harmonic oscillator:

$$x = A\cos(2\pi\nu_1 t) \text{ and } y = B\cos(2\pi\nu_2 t) \tag{4.196}$$

$$p_x = m\dot{x} \text{ and } p_y = m\dot{y} \tag{4.197}$$

The overall motion is not periodic in time unless $\dfrac{\nu_1}{\nu_2}$ is a rational number.
Kepler Problem Example:
Let us do a more extended and detailed example. Returning to the Kepler problem:

$$V(r) = -\frac{k}{r} \tag{4.198}$$

with its separable W:

$$W = W_r(r, \alpha) + W_\theta(\theta, \alpha) + W_\varphi(\varphi, \alpha)\,. \tag{4.199}$$

If we take $E < 0$, we have oscillation in r and θ, along with a rotation in φ. In particular from solving our earlier differential equations for W_θ and W_r, we have

$$\begin{aligned}
W_\varphi &= \alpha_\varphi \varphi \\
W_\theta &= \pm \int d\theta \sqrt{\alpha_\theta^2 - \frac{\alpha_\varphi^2}{\sin^2(\theta)}} \\
W_r &= \pm \int dr \sqrt{2m(\alpha_1 - V(r)) - \frac{\alpha_\theta^2}{r^2}}
\end{aligned}$$

Here we have

$$J_\varphi = \oint p_\varphi \, d\varphi = \oint \frac{\partial W}{\partial \varphi} \, d\varphi = \oint \alpha_\varphi \, d\varphi \tag{4.200}$$

For the cyclic variable φ, we still call the constant p_φ periodic and will take the period to be 2π (arbitrarily since any period would work), which corresponds to particle returning to the original point in space. Thus

$$J_\varphi = 2\pi\alpha_\varphi, \tag{4.201}$$

where α_φ is the angular momentum about $\hat{z}$.
Continuing, in a similar manner we have

$$J_\theta = \oint p_\theta \, d\theta = \oint \frac{\partial W}{\partial \theta} \, d\theta = \oint \pm\sqrt{\alpha_\theta^2 - \frac{\alpha_\varphi^2}{\sin^2(\theta)}} \, d\theta \tag{4.202}$$

Let us call:

$$\cos(\gamma) \equiv \frac{\alpha_\varphi}{\alpha_\theta} \tag{4.203}$$

the angular momentum fraction. Then:

$$J_\theta = \alpha_\theta \oint \pm\sqrt{1 - \frac{\cos^2(\gamma)}{\sin^2(\theta)}}\, d\theta \tag{4.204}$$

If we let $\sin(\theta_0) = \cos(\gamma)$, then $p_\theta = 0$ at the turning points, $\theta \in \{\theta_0, \pi - \theta_0\}$, as expected.

Here one oscillator goes from $\pi - \theta_0 \to \theta_0$ when $p_\theta > 0$, and in reverse for $p_\theta < 0$. Moreover, $\sin(\theta)^{-2}$ is even about $\theta = \frac{\pi}{2}$. This gives

$$J_\theta = 4\alpha_\theta \int_{\frac{\pi}{2}}^{\theta_0} \sqrt{1 - \frac{\cos^2(\gamma)}{\sin^2(\theta)}}\, d\theta. \tag{4.205}$$

Making two more substitutions

$$\cos(\theta) \equiv \sin(\gamma)\sin(\psi)\,, \text{ and then } u \equiv \tan(\psi)\,, \tag{4.206}$$

after some work the expression becomes

$$\begin{aligned} J_\theta &= 4\alpha_\theta \int_0^\infty \left(\frac{1}{1+u^2} - \frac{\cos^2(\gamma)}{1+u^2\cos^2(\gamma)}\right) du = 2\pi\alpha_\theta(1-\cos(\gamma)) \\ &= 2\pi(\alpha_\theta - \alpha_\varphi), \end{aligned} \tag{4.207}$$

which gives

$$J_\theta + J_\varphi = 2\pi\alpha_\theta. \tag{4.208}$$

Finally we can consider

$$J_r = \oint \sqrt{2mE - 2mV(r) - \frac{(J_\theta + J_\varphi)^2}{4\pi^2 r^2}}\, dr \tag{4.209}$$

We can immediately make some observations. We observe that $J_r = J_r(E, J_\theta + J_\varphi)$ is a function of two variables for any $V = V(r)$, and thus if we invert $E = E(J_r, J_\theta + J_\varphi)$. This implies:

$$\frac{\partial E}{\partial J_\theta} = \frac{\partial E}{\partial J_\varphi} \quad \Rightarrow \quad \nu_\theta = \nu_\varphi \tag{4.210}$$

The two frequencies are degenerate for any $V = V(r)$.

For the $V(r) = -kr^{-1}$ potential, the integration can be performed (for example, by contour integration) to give (for $E < 0$)

$$J_r = -(J_\theta + J_\varphi) + \pi k\sqrt{\frac{2m}{-E}} \Rightarrow J_r + J_\theta + J_\varphi = \pi k\sqrt{\frac{2m}{-E}}\,. \tag{4.211}$$

This means:

$$E = -\frac{2\pi^2 k^2 m}{(J_r + J_\theta + J_\varphi)^2} \quad \Rightarrow \quad \nu_\theta = \nu_\varphi = \nu_r \tag{4.212}$$

In particular:

$$\nu_r = \frac{\partial E}{\partial J_r} = 4\pi^2 k^2 (J_r + J_\theta + J_\varphi)^{-3} = \frac{1}{\pi k}\sqrt{\frac{-2E^3}{m}} \tag{4.213}$$

which is the correct orbital frequency in a bound Kepler orbit.

Using the relations between $\{\alpha_1 = E, \alpha_\theta, \alpha_\varphi\}$ and $\{J_r, J_\theta, J_\varphi\}$, we can also get Hamilton's characteristic function for this system as

$$\begin{aligned} W &= W_\varphi + W_\theta + W_r \\ &= \frac{\varphi J_\varphi}{2\pi} \pm \int \sqrt{(J_\theta + J_\varphi)^2 - \frac{J_\varphi^2}{\sin^2(\theta)}}\,\frac{d\theta}{2\pi} \pm \int \sqrt{\frac{-(2\pi m k)^2}{(J_r + J_\theta + J_\varphi)^2} + \frac{2mk}{r} - \frac{(J_\theta + J_\varphi)^2}{(2\pi r)^2}}\, dr. \end{aligned}$$

This then gives the angle variables:

$$\begin{aligned} \omega_r &= \frac{\partial W}{\partial J_r} = \omega_r(r, J_r + J_\theta + J_\varphi, J_\theta + J_\varphi) \\ \omega_\theta &= \frac{\partial W}{\partial J_\theta} = \omega_\theta(r, \theta, J_r + J_\theta + J_\varphi, J_\theta + J_\varphi, J_\varphi) \\ \omega_\varphi &= \frac{\partial W}{\partial J_\varphi} = \omega_\varphi(r, \theta, \varphi, J_r + J_\theta + J_\varphi, J_\theta + J_\varphi, J_\varphi) \end{aligned} \tag{4.214}$$

where $\dot\omega_r = \nu_r$, $\dot\omega_\theta = \nu_\theta$, and $\dot\omega_\varphi = \nu_\varphi$. Of course, in this case, $\nu_r = \nu_\theta = \nu_\varphi$.

At this point we can identify five constants for the Kepler problem

$$\begin{aligned} J_1 &= J_\varphi \\ J_2 &= J_\theta + J_\varphi \\ J_3 &= J_r + J_\theta + J_\varphi \\ \omega_1 &= \omega_\varphi - \omega_\theta \\ \omega_2 &= \omega_\theta - \omega_r\,. \end{aligned} \tag{4.215}$$

(These 5 constants could also be identified from the angular momentum $\vec{L}$, energy E, and Laplace-Runge-Lenz vector $\vec{A}$.) What are they? There are two constants specifying the plane of the orbit (the $x'y'$-plane), which are the inclination angle i and the longitude of the ascending node Ω. There are three constants specifying the form of the ellipse, which are the semi-major axis a (giving the size), the eccentricity ε (giving the shape), and the angle ω (giving the orientation within the plane). These are all shown in Fig. 4.9.

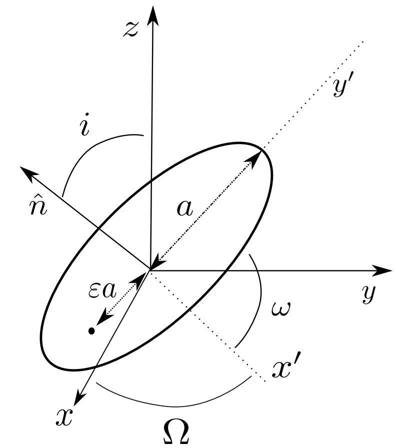


Figure 4.9: Picture of an orbit in 3d and the five parameters necessary to fully specify it. The angles i, Ω and ω provide the orientation in space while a and ε provide the size and shape of the conic section.

It can be shown that the relations between these constants and the ones above are

$$\cos(i) = \frac{J_1}{J_2} \qquad a = -\frac{k}{2E} = \frac{J_3^2}{4\pi^2 mk} \qquad \epsilon = \sqrt{1 - \left(\frac{J_2}{J_3}\right)^2}$$

$$\Omega = 2\pi\omega_1 \qquad \omega = 2\pi\omega_2$$

providing a fairly simple physical interpretations to $(J_1, J_2, J_3, \omega_1, \omega_2)$. Also recall that $J_2 = 2\pi\alpha_\theta = 2\pi\ell$.

When the orbit is perturbed by additional forces (like a moon, another planet, general relativistic corrections, and so on), these action-angle variables provide a natural way to describe the modified orbit. We will soon see that they become functions of time. For example, from the above list of constants, the perturbed $\omega = \omega(t)$ is the precession of the perihelion of an orbit. We will learn how to derive this time dependence next, in our chapter on Perturbation Theory.