Chapter 1

A Review of Analytical Mechanics

1.1 Introduction

These lecture notes cover the third course in Classical Mechanics, taught at MIT since the Fall of 2012 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; generalized coordinates could be angles, et cetera.

Velocities are 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 $\mathbf{p}_i = m_i \mathbf{v}_i$ or generalized 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*. These are frames where the motion of a particle not subject to forces is in a straight line with constant velocity. The converse does not hold. Inertial frames describe time and space homogeneously (invariant to displacements), isotropically (invariant to rotations), and in a time independent manner. 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 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 a strange manner when going between different frames).

We will use primes will denote coordinate transformations. If $\mathbf{r}$ is measured in an inertial frame S, and $\mathbf{r}'$ is measured in frame S' with relation to S 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,

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

which preserves the inertiality of frames, with $\mathbf{F} = m\ddot{\mathbf{r}}$ and $\mathbf{F}' = m\ddot{\mathbf{r}}'$ implying each other. 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 $\ddot{\mathbf{r}} = 0$ still gives $\ddot{\mathbf{r}}' \neq 0$, so the primed frame is noninertial.

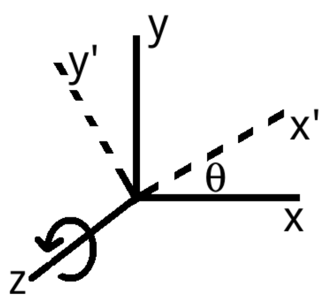

Figure 1.1: Frame rotated by an angle θ

2. In polar coordinates, $\mathbf{r} = r\hat{r}$, gives

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

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, $q = (q_1, \dots, q_N)$ and likewise $\dot{q} = (\dot{q}_1, \dots, \dot{q}_N)$. We are now letting N denote the number of scalar (rather than vector) variables, and will often use the short form to denote dependence on these variables, as in Eq. (1.3). Typically we can write $L = T - V$ where T is the kinetic energy and V is the potential energy. In the simplest cases, $T = T(\dot{q})$ and $V = V(q)$, but we also allow the more general possibility that $T = T(q, \dot{q}, t)$ and $V = V(q, \dot{q}, t)$. It turns out, as we will discuss later, that even this generalization does not describe all possible classical mechanics problems.

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

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

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

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

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

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

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).

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 $q_2 = \theta$.
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 $T = \frac{m}{2}\dot{\mathbf{r}}^2 = \frac{m}{2}\left(\dot{r}^2 + r^2\dot{\theta}^2\right)$, giving

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

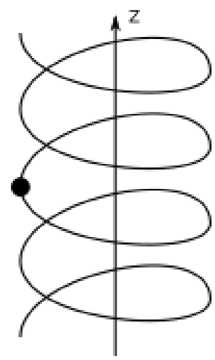

Figure 1.2: Bead on a helix

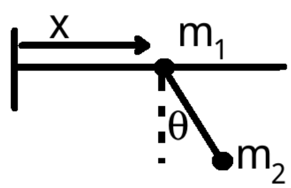

Figure 1.3: Pendulum of mass m_2 hanging on a rigid bar of length ℓ whose support m_1 is a frictionless horizontally sliding bead

For r the Euler-Lagrange equation is

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

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

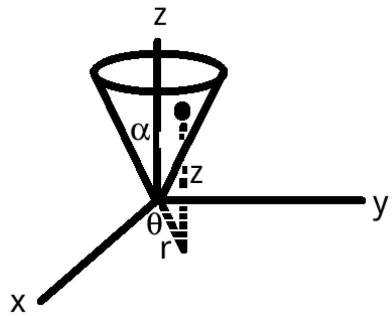

Figure 1.4: Particle on the inside of a cone

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
$$\left(1+\cot^2\alpha\right)\ddot{r} - r\dot{\theta}^2 + g\cot\alpha = 0. \tag{1.21}$$
For θ we have $0 = \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{\theta}}\right) - \frac{\partial L}{\partial \theta}$, so
$$0 = \frac{d}{dt}\left(mr^2\dot{\theta}\right) - 0, \tag{1.22}$$
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.

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 $2N$ 1st 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

$$dH = \frac{\partial H}{\partial q} dq + \frac{\partial H}{\partial p} dp + \frac{\partial H}{\partial t} dt \tag{1.27}$$

$$= \cancel{p d\dot{q}} + \dot{q} dp - \frac{\partial L}{\partial q} dq - \cancel{\frac{\partial L}{\partial \dot{q}} d\dot{q}} - \frac{\partial L}{\partial t} dt\,. \tag{1.28}$$

Since we are free to independently vary dq, dp, and dt this implies $\frac{\partial L}{\partial \dot{q}} = p$, $\frac{\partial L}{\partial q} = \dot{p}$, and $\frac{\partial H}{\partial t} = -\frac{\partial L}{\partial t}$.

We can interpret the two Hamilton equations as follows:

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

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

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

$$\dot q_i p_i = \dot q_i \frac{\partial L}{\partial \dot q_i} = \frac{1}{2}\dot q_i T ik \dot q_k + \frac{1}{2}\dot q_j T_{ji}\dot q_i = 2T. \tag{1.30}$$

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

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

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

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

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

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

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):

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

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

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

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

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

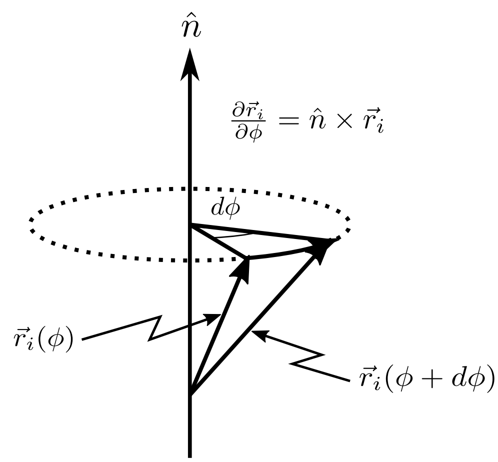

Figure 1.5: Rotation about a symmetry axis

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,

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

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. Here the scaling for time is given by $1 - k/2 = 0$, so it does not change 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 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

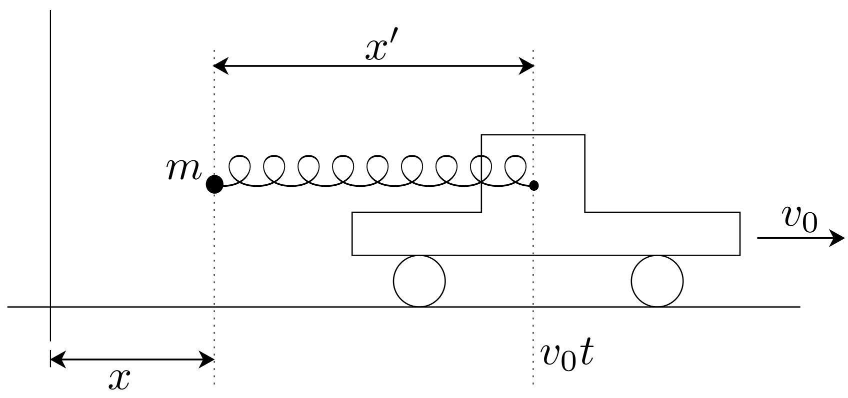

Figure 1.6: Mass attached by a spring to a moving cart

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

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

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

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

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

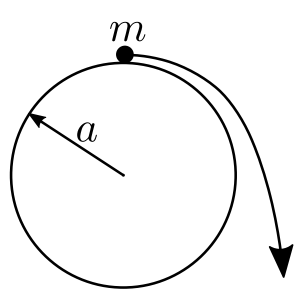

Figure 1.7: Mass on a sphere

  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*.

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 remain parallel to z, so the

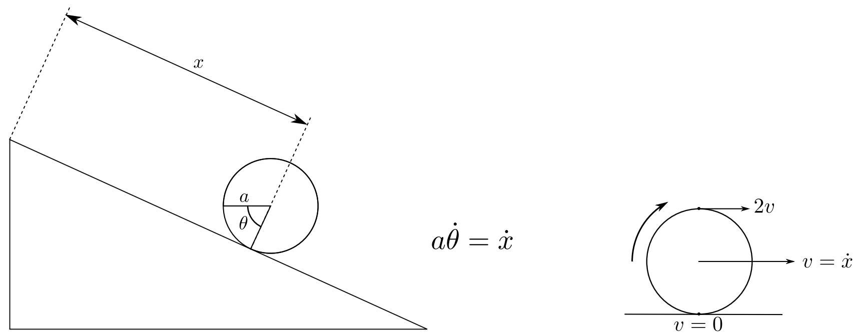

Figure 1.8: Disk rolling down an incline without slipping

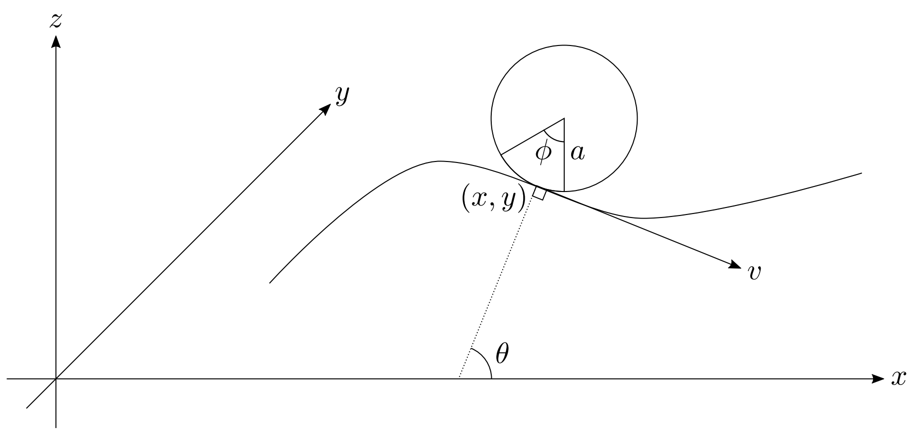

Figure 1.9: Vertical rolling disk on a two dimensional plane

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*, following the terminology of Goldstein.

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

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

so $a_j = \frac{\partial f}{\partial q_j}$ and $a_t = \frac{\partial f}{\partial t}$. The symmetry 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 $\frac{\partial h}{\partial \theta} = 0$ which is a contradiction for a 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

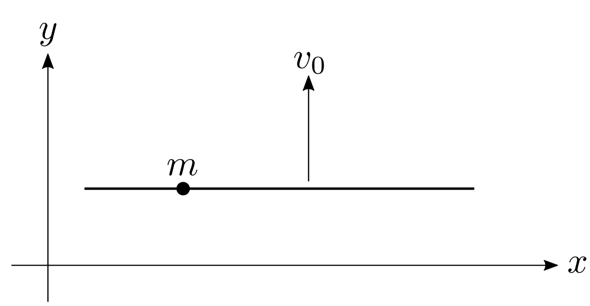

Figure 1.10: Bead on a moving wire

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.

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

$$\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_j}\right) - \frac{\partial T}{\partial q_j} - Q_j = 0\,. \tag{1.69}$$

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

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

   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.

An important example of a nonconservative forces R_j is given by *friction*.

- Static friction is $F_{\rm s} \le F_{\rm s}^{\rm max} = \mu_s F_{\rm N}$ for a normal force $F_{\rm N}$.
- Sliding friction is $\mathbf{F} = -\mu F_{\rm N}\,\frac{\mathbf{v}}{v}$, so this is a constant force that is always opposite the direction of motion (but vanishes when there is no motion).
- Rolling friction is $\mathbf{F} = -\mu_{\rm R} F_{\rm N}\,\frac{\mathbf{v}}{v}$.
- 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 right-hand-side). For this form

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

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 $\dot{y}_H = A\exp(-6\pi\eta a t/m')$, where the constant A must be determined by an initial condition. The result $\dot{y} = \dot{y}_H + \dot{y}_p$ can be integrated in time once more to obtain the full solution $y(t)$ for the motion.

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

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

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 in lecture.

1.5 Calculus of Variations & Lagrange Multipliers

Calculus of Variations

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

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

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^{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 $\int_{s_1}^{s_2} \sum_i \left[\frac{\partial f}{\partial y_i} - \frac{d}{ds}\left(\frac{\partial f}{\partial \dot{y}_i}\right)\right] \delta y_i(s)\, ds = 0$. For independent variations δy_i (for example, after imposing holonomic constraints), this can only occur if

$$\frac{\partial f}{\partial y_i} - \frac{d}{ds}\left(\frac{\partial f}{\partial \dot{y}_i}\right) = 0\,. \tag{1.78}$$

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. (1.78).

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,

$$ds = \sqrt{(dx)^2 + (dy)^2 + (dz)^2} = \sqrt{a^2(d\theta)^2 + a^2\sin^2(\theta)(d\phi)^2} \tag{1.79}$$

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

$$ds = a\sqrt{1 + \sin^2(\theta)\left(\frac{d\phi}{d\theta}\right)^2}\, d\theta, \tag{1.80}$$

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

$$\delta J[q,p] = \delta \int_{t_1}^{t_2} dt \left[p_i \dot{q}_i - H(q,p,t) \right] = 0 \tag{1.81}$$

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, of 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

$$\begin{aligned} \frac{d}{dt}\left(\frac{\partial f}{\partial \dot{q}_i}\right) - \frac{\partial f}{\partial q_i} = 0 \qquad &\Longrightarrow \qquad \dot{p}_i = -\frac{\partial H}{\partial q_i}\,, \\ \frac{d}{dt}\left(\frac{\partial f}{\partial \dot{p}_i}\right) - \frac{\partial f}{\partial p_i} = 0 \qquad &\Longrightarrow \qquad \dot{q}_i = \frac{\partial H}{\partial p_i}\,, \end{aligned} \tag{1.82}$$

giving Hamilton's equations as expected. Note that because f is independent of $\dot{p}_i$, the term $(\partial f/\partial \dot{p}_i)\delta p_i = 0$, 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 (which we will exploit later in detail when discussing canonical transformations).

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

$$\left(\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_j}\right) - \frac{\partial L}{\partial q_j} - R_j\right)\delta q_j = 0 \tag{1.83}$$

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

$$\delta J = \int_{t_1}^{t_2} dt \left(\frac{\partial L}{\partial q_j} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_j}\right)\right)\delta q_j = 0 \tag{1.84}$$

(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. 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

$$g_\alpha(q, \dot{q}, t) = a_{j\alpha}(q,t)\dot{q}_j + a_{t\alpha}(q,t) = 0 \tag{1.85}$$

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

$$\sum_{j=1}^{n} a_{j\alpha} \delta q_j = 0\,, \tag{1.86}$$

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 yields

$$\left[\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_j}\right) - \frac{\partial L}{\partial q_j} - \lambda_\alpha a_{j\alpha}\right] \delta q_j = 0 \tag{1.87}$$

where the sums implicitly run over both α and j. Its 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:

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

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 (1.88) in conjunction with $g_\alpha = 0$ for each α from (1.85) 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

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_j}\right) - \frac{\partial L}{\partial q_j} = \sum_{\alpha=1}^{k} \lambda_\alpha \frac{\partial f_\alpha}{\partial q_j} \tag{1.89}$$

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

$$J[q_j, \lambda_\alpha] = \int_{t_1}^{t_2} \left(L + \lambda_\alpha f_\alpha\right) dt \tag{1.90}$$

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 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
$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_j}\right) - \frac{\partial L}{\partial q_j} = \sum_{\alpha=1}^{k} \lambda_\alpha \frac{\partial g_\alpha}{\partial \dot{q}_j} \tag{1.91}$$
for semi-holonomic constraints. This result cannot be derived from Hamilton's principle in general, justifying the time we spent discussing d'Alembert's principle, which we have used to obtain (1.91). 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, $\frac{\partial \mathcal{F}}{\partial \dot{q}}$ from 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
$$m\ddot{r} - mr\dot{\theta}^2 + mg\cos(\theta) = \lambda, \tag{1.92}$$
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
$$\frac{d}{dt}\left(mr^2\dot{\theta}\right) - mgr\sin(\theta) = 0. \tag{1.93}$$
This in conjunction with $r = a$ gives 3 equations for the 3 variables (r, θ, λ). Putting them together 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 $\dot{\theta}^2 = \frac{2g}{a}(1 - \cos(\theta))$. Thus,
$$\lambda = mg(3\cos(\theta) - 2) \tag{1.94}$$

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

$$L = \frac{m}{2}\left(\dot{x}^2 + \dot{z}^2\right) - mgz, \tag{1.95}$$

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 simply left as $\dot{z}_1$ and $\dot{x}_1$ for simplicity in writing (though the actual values follow from $\dot{z}_1 = -a\sin\theta_1\dot{\theta}_1$ and $\dot{x}_1 = a\cos\theta_1\dot{\theta}_1$). This means

$$x(t) = \dot{x}_1(t - t_1) + x_1, \tag{1.96}$$
$$z(t) = -\frac{g}{2}(t - t_1)^2 + \dot{z}_1(t - t_1) + z_1, \tag{1.97}$$

where t_1 is the time when the mass leaves the sphere. That can be found from

$$\dot{\theta}^2 = \frac{2g}{a}(1 - \cos(\theta)) = \frac{4g}{a}\sin^2\left(\frac{\theta}{2}\right), \tag{1.98}$$

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.

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 $h = a\dot{\theta} - \dot{x} = a$, which can be made holonomic but which we will treat as semi-holonomic. Then $T = T_{\text{CM}} + T_{\text{rotation}} = \frac{1}{2}m\dot{x}^2 + \frac{1}{2}ma^2\dot{\theta}^2$ as $I_{\text{hoop}} = ma^2$. Meanwhile, $V = mg(l - x)\sin(\phi)$ so that $V(x = l) = 0$. This means

$$L = T - V = \frac{m}{2}\dot{x}^2 + \frac{ma^2}{2}\dot{\theta}^2 + mg(x - l)\sin(\phi). \tag{1.99}$$

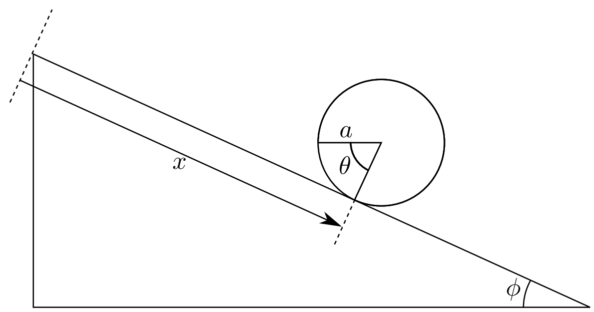

Figure 1.11: Hoop rolling on inclined plane

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

$$m\ddot{x} - mg\sin(\phi) = \lambda \quad \text{and} \quad ma^2\ddot{\theta} = \lambda a, \tag{1.100}$$

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

$$\lambda = \frac{1}{2} mg\sin(\phi) \tag{1.101}$$

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

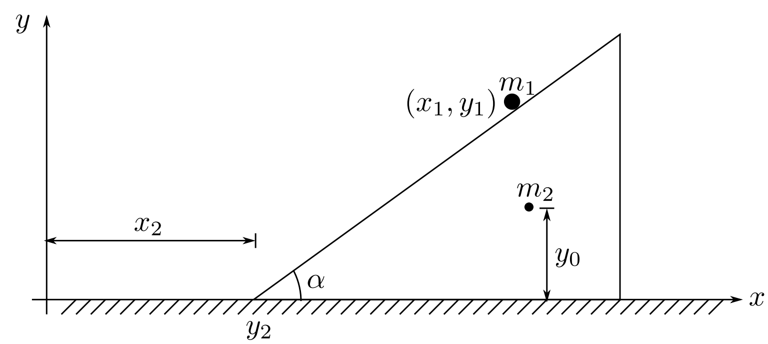

Figure 1.12: Wedge sliding on ice

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 only ask about these two.) 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 CM of the wedge taken from above its bottom. 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 holonomic constraints are

$$\begin{aligned}
\frac{d}{dt}\frac{\partial L}{\partial \dot{x}_1} - \frac{\partial L}{\partial x_1} = \lambda_1 \frac{\partial f_1}{\partial x_1} + \lambda_2 \frac{\partial f_2}{\partial x_1} \qquad &\Longrightarrow \qquad m_1 \ddot{x}_1 = -\lambda_1 \tan(\alpha), \\
\frac{d}{dt}\frac{\partial L}{\partial \dot{y}_1} - \frac{\partial L}{\partial y_1} = \lambda_1 \frac{\partial f_1}{\partial y_1} + \lambda_2 \frac{\partial f_2}{\partial y_1} \qquad &\Longrightarrow \qquad m_1 \ddot{y}_1 + m_1 g = \lambda_1, \\
\frac{d}{dt}\frac{\partial L}{\partial \dot{x}_2} - \frac{\partial L}{\partial x_2} = \lambda_1 \frac{\partial f_1}{\partial x_2} + \lambda_2 \frac{\partial f_2}{\partial x_2} \qquad &\Longrightarrow \qquad m_2 \ddot{x}_2 = \lambda_1 \tan(\alpha), \\
\frac{d}{dt}\frac{\partial L}{\partial \dot{y}_2} - \frac{\partial L}{\partial y_2} = \lambda_1 \frac{\partial f_1}{\partial y_2} + \lambda_2 \frac{\partial f_2}{\partial y_2} \qquad &\Longrightarrow \qquad m_2 \ddot{y}_2 + m_2 g = -\lambda_1 + \lambda_2,
\end{aligned} \tag{1.102}$$

which in conjunction with $y_1 - y_2 = (x_1 - x_2)\tan(\alpha)$ and $y_2 = 0$ is six equations. We number them (1) to (6). 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

$$\lambda_1 = \frac{g}{\frac{1}{m_1 \cos^2(\alpha)} + \frac{\tan^2(\alpha)}{m_2}}. \tag{1.103}$$

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

$$\begin{aligned}
\ddot{x}_2 &= \frac{\tan(\alpha)}{m_2}\lambda_1, \\
\ddot{x}_1 &= -\frac{\tan(\alpha)}{m_1}\lambda_1, \\
\ddot{y}_1 &= \frac{\lambda_1}{m_1} - g,
\end{aligned} \tag{1.104}$$

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).