8.223 Classical Mechanics II

MIT, IAP 2018

Reuben R.W. Wang

Contents

Chapter 1

Introduction

Here, we will be introduced to a new way of viewing classical mechanics. How we approach solving problems and the techniques we employ to do so will be a new step away from just $\vec{F} = m\vec{a}$. But before we begin formulating these new tools, we first do a little recap on the *Newtonian mechanics* method. There is a systematic approach in attempting all Newtonian mechanics problems. The steps are listed as follows:

1. Set-up the coordinate system.
2. Draw the free body diagram.
3. Write the equations of motion.
4. Eliminate the forces of constraint.
5. Solve for the final equations of motion.

To illustrate this, we shall work out the following simple example.

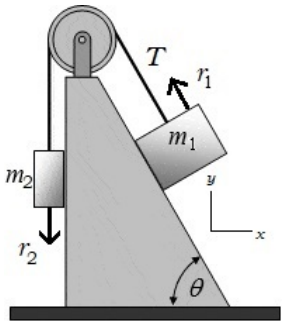

Figure 1.1: 2 Mass Ramp-Pulley System

Consider a 2 mass ramp-(massless)pulley system as shown in figure 1.1. We will first use the steps listed above to solve this problem.

1. Setting up a suitable coordinate system, we can define the following quantities

$$r_1 = r_2 = r \tag{1.1}$$
$$x_1 = x_0 \tag{1.2}$$
$$y_1 = y_0 - r \tag{1.3}$$
$$x_2 = r\cos(\theta) \tag{1.4}$$
$$y_2 = r\sin(\theta) \tag{1.5}$$

   And the respective unit vectors are

$$\hat{r_1} = \cos(\theta)\hat{x} + \sin(\theta)\hat{y} \tag{1.6}$$
$$\hat{r_2} = -\hat{y} \tag{1.7}$$
$$\hat{\theta} = -\sin(\theta)\hat{x} + \cos(\theta)\hat{y} \tag{1.8}$$

2. Next, we draw the free-body diagrams of m_1 and m_2 as seen in figures 1.3 and 1.2. Notice that m_2 is constrained to only move in the y direction.

3. Utilizing the free body diagrams, we set-up the respective equations of motion for both masses.

$$F_{x_1} = T\cos(\theta) - N\sin(\theta) \tag{1.9}$$
$$F_{y_1} = T\sin(\theta) + N\cos(\theta) - m_1 g \tag{1.10}$$
$$F_{x_2} = 0 \tag{1.11}$$
$$F_{y_2} = T - m_2 g \tag{1.12}$$

4. Looking at the set-up, we formulate the forces of constraint as follows,

$$\dot{\theta} = 0 \tag{1.13}$$
$$\Rightarrow \quad \vec{F_1}\cdot\hat{\theta} = \vec{0} \tag{1.14}$$
$$\Rightarrow \quad N = m_1 g\cos(\theta) \tag{1.15}$$
$$r_1 = r_2 \quad \Rightarrow \quad \dot{r_1} = \dot{r_2} \quad \Rightarrow \quad \ddot{r_1} = \ddot{r_2} \tag{1.16}$$
$$\Rightarrow \quad \frac{\vec{F_1}\cdot\hat{r}_1}{m_1} = \frac{T}{m_1} - g\sin(\theta) = \frac{\vec{F_2}\cdot\hat{r}_2}{m_2} = -\frac{T}{m_2} + g$$
$$\Rightarrow \quad T\big(\frac{1}{m_1} + \frac{1}{m_2}\big) = g\big(1+\sin(\theta)\big)$$
$$\Rightarrow \quad \boxed{T = \big(\frac{m_1 m_2}{m_1+m_2}\big)g\big(1+\sin(\theta)\big)} \tag{1.17}$$

5. Then finally, writing out the final equation of motion

$$\ddot{r} = -\frac{F_{y_2}}{m_2} = g - \frac{T}{m_2}$$
$$= g - \big(\frac{m_1 m_2}{m_1+m_2}\big)\frac{g}{m_2}\big(1+\sin(\theta)\big)$$
$$\Rightarrow \quad \boxed{\ddot{r} = \big(\frac{g}{m_1+m_2}\big)\big(m_2 - m_1\sin(\theta)\big) = \text{constant}} \tag{1.18}$$

Note that dots above the dynamical variables represent time derivatives.

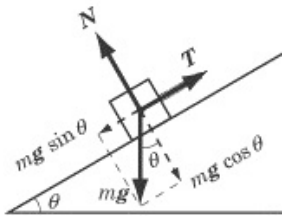

Figure 1.2: m_1 Free Body Diagram

Figure 1.3: m_2 Free Body Diagram

§1.1 Generalizing $\vec{F} = m\vec{a}$

Let's now try a different approach to solve the example above by looking a little deeper into Newton's famous equation,

$$F = -\frac{\partial V}{\partial x} = m\ddot{x} \tag{1.19}$$
$$= \frac{d}{dt}(m\dot{x}) = \dot{p}$$
$$m\dot{x} = \frac{\partial}{\partial \dot{x}}\left(\frac{1}{2}m\dot{x}^2\right) = \frac{\partial T}{\partial \dot{x}} \tag{1.20}$$
$$\Rightarrow \boxed{-\frac{\partial V}{\partial x} = \frac{d}{dt}\frac{\partial T}{\partial \dot{x}}} \tag{1.21}$$

Where T denotes the kinetic energy of the system (**not** to be confused with string tension).

(1.21) is known as the *generalized* form of Newton's force law and is going to greatly simplify the required steps for the 2 mass ramp-pulley problem as we will show below.

$$V = mgy_1 + mgy_2 = g(m_2\sin(\theta) - m_2)r \tag{1.22}$$
$$T = \frac{1}{2}\left(m_1v_1^2 + m_2v_2^2\right) = \frac{1}{2}(m_1 + m_2)\dot{r}^2 \tag{1.23}$$

since

$$\dot{x} = \dot{r}\cos(\theta)$$
$$\Rightarrow \frac{\partial T}{\partial \dot{r}} = (m_1 + m_2)\dot{r}$$
$$-\frac{\partial V}{\partial r} = g\left(m_2 - m_1\sin(\theta)\right)$$

then substituting these results into (1.21), we get

$$\ddot{r} = \left(\frac{g}{m_1 + m_2}\right)\left(m_2 - m_1\sin(\theta)\right) \tag{1.24}$$

Which is the same result in considerably less steps!

Chapter 2

Lagrangian Mechanics

We now embark on the formal study of *Lagrangian mechanics*, which is a reformulation of classical mechanics by Joseph-Louis Lagrange in 1788. We will focus on the *Lagrange equation* of the second kind in these notes, hence we must first begin with the notion of *generalized coordinates*.

§2.1 Generalized Coordinates

When we think about coordinates, it is likely our mind jumps to Cartesian coordinates, or perhaps spherical or cylindrical coordinate systems as well. However for the formalism of classical mechanics we're about to construct, it will be useful to have a general system of coordinates that can take on the most convenient set of variables to describe a system. How do we know the number of such generalized coordinates to have? We look to what is known as the *degrees of freedom* of the system.

Definition 2.1.1. Degrees of Freedom: *For a given mechanical system, its degrees of freedom are the number of independent parameters required to uniquely define its configuration.*

For a dynamical system with N degrees of freedom, we require N generalized coordinates to describe them. We label these coordinates as

$$\{q_j\}, \quad j = 1, 2, ..., N \tag{2.1}$$

Where the set of first derivatives with respect to time of these coordinates are know as the *generalized velocities* of the system labelled as

$$\{\dot{q}_j(t)\} = \{\frac{dq_j(t)}{dt}\} \tag{2.2}$$

With these, we can then completely describe the state of a dynamical system due to the following postulate,

Postulate 2.1.1. *The system state is* ***completely*** *specified by* $\{q_j\}$ *and* $\{\dot{q}_j\}$ *provided we know the initial conditions* $q_j(t = 0)$ *and* $\dot{q}_j(t = 0)$ $\forall j$.

With this new problem, the following example will show that (1.21) is actually no longer sufficient for it to be applied to any arbitrary system (because (1.21) implicitly assumes the use of independent coordinates). We will show this in the following example.

Example

Consider again the system as shown in figure 1.1 but now, m_2 is allowed to swing freely in a pendulum-like fashion. We then relabel the angle of inclination of the ramp to be θ_0 and the angle of m_2 from the vertical to be θ instead. We also take the $\hat{\theta}$ unit vector of m_2 to be pointing anti-clockwise.

Our system now has 2 degrees of freedom,

1. $r_1 = r_2 = r$
2. θ

Before we implement the use of (1.21) to show that it fails, we first use the tried and tested Newtonian method to attain the correct analytical solution. We use basically the same coordinate set-up as in the first example but noting the appropriate change from θ to θ_0. We also define new local Cartesian coordinates for m_2 a label them with primes to differentiate these from the global x, y.

As such, the free body diagram for m_2 is now modified to

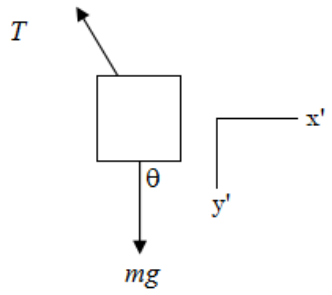

Figure 2.1: m_2 Free Body Diagram

Notice here that we have defined the y' coordinate being downward as positive for convenience. The equations of motion for m_2 are thus

$$m_2\ddot{x}_2 = -T\sin(\theta) \tag{2.3}$$
$$m_2\ddot{y}_2 = m_2 g - T\cos(\theta) \tag{2.4}$$

Also from the geometry of the problem, we see that

$$\ddot{y}_2 = r_2 \cos(\theta) \tag{2.5}$$

$$\Rightarrow \ddot{y}_2 = \frac{d}{dt}\big(\dot{r}\cos(\theta) - r_2\dot{\theta}\sin(\theta)\big)$$

$$\Rightarrow \ddot{y}_2 = \ddot{r}\cos(\theta) - 2\dot{r}_2\dot{\theta}\sin(\theta) - r_2\big(\ddot{\theta}\sin(\theta) + \dot{\theta}^2\cos(\theta)\big)$$

$$\Rightarrow g - \frac{T}{m_2}\cos(\theta) = \ddot{r}\cos(\theta) - 2\dot{r}_2\dot{\theta}\sin(\theta) - r_2\big(\ddot{\theta}\sin(\theta) + \dot{\theta}^2\cos(\theta)\big) \tag{2.6}$$

$$x_2 = r_2\sin(\theta) \tag{2.7}$$

$$\Rightarrow \ddot{x}_2 = \ddot{r}_2\sin(\theta) + 2\dot{r}_2\dot{\theta}\cos(\theta) + r_2\big(\ddot{\theta}\cos(\theta) - \dot{\theta}^2\sin(\theta)\big)$$

$$\Rightarrow -\frac{T}{m_2}\sin(\theta) = \ddot{r}_2\sin(\theta) + 2\dot{r}_2\dot{\theta}\cos(\theta) + r_2\big(\ddot{\theta}\cos(\theta) - \dot{\theta}^2\sin(\theta)\big) \tag{2.8}$$

And also from the free body diagram in figure 1.2, we see that

$$m_1\ddot{r}_1 = T - m_1 g\sin(\theta_0) \tag{2.9}$$

Now from the intermediate equations of motion (2.6), (2.8) and (2.9), we end up with these final equations of motion.

$$\ddot{r} = \Big(\frac{g}{m_1+m_2}\Big)\big(m_2\cos(\theta) - m_1\sin(\theta_0)\big) + \Big(\frac{m_2}{m_1+m_2}\Big)r\dot{\theta}^2 \tag{2.10}$$

$$\ddot{\theta}_2 = -\frac{1}{r}\big(g\sin(\theta) + 2\dot{r}\dot{\theta}\big) \tag{2.11}$$

That was horribly tedious, considering we even skipped over a few steps at the end to arrive at the final (2.10) and (2.11) equations of motion.

Let's go ahead and try this problem again using the generalized equations of motion. Remember again that although this will be faster, here we are showing (1.21) is not sufficient in its formulation for all mechanics problems (answer **will** turn out wrong). First, we further generalize (1.21) such that it utilizes generalized coordinates. We do a simple substitution of the variables of differentiation to give

$$-\frac{\partial V}{\partial q_j} = \frac{d}{dt}\frac{\partial T}{\partial \dot{q}_j} \tag{2.12}$$

We have included the subscript indices here to indicate that this is the generalized Newton's law for the j^{th} parameter.

Looking back at our original no-pendulum problem, we kind of implicitly came to the conclusion that this was a problem with 1 degree of freedom r after solving it. But since we need to establish the number of degrees of freedom from the start when using generalized coordinates, how can we do such an analysis? This is best shown from an example, and we will pick our initial no-pendulum problem for this.

1. Our system had 2 masses and was a 2 dimensional problem
   $\Rightarrow$ there were $2 \times 2 = 4$ degrees of freedom.

2. Since we incorporated 2 constraints into the system (string and ramp),
   $\Rightarrow$ This eliminated 2 degrees of freedom leaving us with 2.
3. Finally, we know that m_2 would only be lowered and raised by the pulley ($F_{x_2} = 0$,
   $\Rightarrow$ This eliminates one more degree of freedom leaving us with just the 1, r.

For the second free-pendulum problem, we again initially just stated that there were 2 degrees of freedom without any explicitly justification. But following a similar analysis to what we have done above, we see that it is just step 3 we remove leaving us with exactly 2 degrees of freedom as expected ($q_1 = r$, $q_2 = \theta$).

The next thing to do is then to find the potential and kinetic energies of the system for us to utilize (2.12).

$$V = \sum_j m_j g \triangle h_j \tag{2.13}$$
$$= g\big(m_1 \sin(\theta_0) - m_2 \cos(\theta)\big) r \tag{2.14}$$
$$T = \frac{1}{2} \sum_j m_j v_j^2 \tag{2.15}$$
$$= \frac{1}{2}(m_1 + m_2)\dot{r}^2 + \frac{1}{2} m_2 (r\dot{\theta})^2 \tag{2.16}$$

Where the $\triangle h_j$ terms were found by geometrical means, and the kinetic energy terms are separately the linear and rotational kinetic energies.

$$r : \begin{cases} \frac{\partial V}{\partial r} = g\big(m_1 \sin(\theta_0) - m_2 \cos(\theta)\big) \\ \frac{\partial T}{\partial \dot{r}} = (m_1 + m_2)\dot{r} \end{cases}$$
$$\theta : \begin{cases} \frac{\partial V}{\partial \theta} = g m_2 \sin(\theta) \\ \frac{\partial T}{\partial \dot{\theta}} = m_2 r^2 \dot{\theta} \end{cases}$$

Substituting these into (2.12), we get

r :

$$-g\big(m_1 \sin(\theta_0) - m_2 \cos(\theta)\big) = \frac{d}{dt}(m_1 + m_2)\dot{r}$$
$$\Rightarrow \quad \ddot{r} = \frac{g}{(m_1 + m_2)}\big(m_2 \cos(\theta) - m_1 \sin(\theta_0)\big) \tag{2.17}$$

θ :

$$-g m_2 \sin(\theta) = \frac{d}{dt} m_2 r^2 \dot{\theta}$$
$$\Rightarrow \quad -g \sin(\theta) = 2r\dot{r}\dot{\theta} + r^2 \ddot{\theta}$$
$$\Rightarrow \quad \ddot{\theta} = -\frac{1}{r}(g \sin(\theta) + 2\dot{r}\dot{\theta}) \tag{2.18}$$

Here we see that although the equation of motion for θ agrees with (2.11), the equation of motion for r is **missing** the centripetal term in (2.10)!
What went wrong here? The answer lies in the fact that we had naively assumed $V = V(q)$

and $T = T(\dot{q})$ for our chosen coordinate system. In actual fact, our kinetic energy depends on r! It seems at this point that we need an even further generalization of (1.21).

§2.2 Euler-Lagrange Equations (First Look)

The current aim is to modify (2.12) into an equation that ensures our 'energy' equivalent terms include both q and $\dot{q}$ as function arguments. As such, let us define

$$\mathcal{L}(q, \dot{q}) \equiv T(\dot{q}) - V(q) \tag{2.19}$$

And by our original assumption that $V = V(q)$ and $T = T(\dot{q})$,

$$\begin{aligned} & \frac{\partial V}{\partial \dot{q}} = 0, \quad \frac{\partial T}{\partial q} = 0 \\ \Rightarrow \quad & \frac{\partial \mathcal{L}}{\partial q_j} = \frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{q}_j} \end{aligned} \tag{2.20}$$

This is kind of a guess for a fix, but let us apply it to our ramp-pendulum problem and see if things work out. Notice that we have included indices in (2.20), implying that they run over all the degrees of freedom

Example Continued...

$$\begin{aligned} \mathcal{L} &= T - V \\ &= \frac{1}{2}(m_1 + m_2)\dot{r}^2 + \frac{1}{2}m_2(r\dot{\theta})^2 - g\big(m_1 \sin(\theta_0) - m_2 \cos(\theta)\big)r \\ \Rightarrow \quad \frac{\partial \mathcal{L}}{\partial r} &= m_2 r\dot{\theta}^2 - g\big(m_1 \sin(\theta_0) - m_2 \cos(\theta)\big) \qquad (2.21) \\ \frac{\partial \mathcal{L}}{\partial \dot{r}} &= (m_1 + m_2)\dot{r} \qquad (2.22) \end{aligned}$$

Then substituting these into (2.20), we get

$$\begin{aligned} \frac{d}{dt}(m_1 + m_2)\dot{r} &= m_2 r\dot{\theta}^2 - g\big(m_1 \sin(\theta_0) - m_2 \cos(\theta)\big) \\ \Rightarrow \quad \ddot{r} &= \frac{g}{(m_1 + m_2)}\big(m_2 \cos(\theta) - m_1 \sin(\theta_0)\big) + \frac{m}{(m_1 + m_2)_2} r\dot{\theta}^2 \qquad (2.23) \end{aligned}$$

Amazingly, this matches perfectly equation (2.10)! Just to be certain, we also work out the equations of motion for θ.

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial \theta} &= -g m_2 r \cos(\theta) \qquad (2.24) \\ \frac{\partial \mathcal{L}}{\partial \dot{\theta}} &= m_2 r^2 \dot{\theta} \qquad (2.25) \\ \Rightarrow \quad m_2 r^2 \ddot{\theta} + 2m_2 r\dot{r}\dot{\theta} &= -g m_2 r \sin(\theta) \\ \Rightarrow \quad \ddot{\theta} &= -\frac{1}{r}\big(g \sin(\theta) + 2\dot{r}\dot{\theta}\big) \qquad (2.26) \end{aligned}$$

which is exactly what we got before.

In actual fact, this 'generalized' equation is known as the *Euler-Lagrange* equation and is the governing equation for Lagrangian mechanics. Just to formally restate it

$$\boxed{\frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{q}_j} - \frac{\partial \mathcal{L}}{\partial q_j} = 0} \tag{2.27}$$

If you found the derivation of (2.27) a little hand wavey, that's because it was! We will move on to show a different means of arriving at this result in a more concrete manner.

§2.3 Principle of Least Action

The *principle of least action* is an idea that loosely speaking, tells us how trajectories will always take the 'easiest' path between 2 points. This principal is what we will be using to better derive (2.27) and as such, we will present it formally.

Principal of Least Action

For some function $\mathcal{L} = \mathcal{L}(q, \dot{q}, t)$, the motion of the system always ***minimizes*** *the 'action' S, defined as*

$$S = \int_{t_1}^{t_2} \mathcal{L}(q, \dot{q}, t)dt \tag{2.28}$$

for given start and end points $q(t_1)$ and $q(t_2)$ respectively.

At this point in our formalizing of this principal, we take it that we do not yet know $\mathcal{L}$ is the Lagrangian. Hence we now need to derive $\mathcal{L}$ such that it solves this problem of minimizing the integral above.

§2.4 Euler-Lagrange Equations (Second Look)

For simplicity, we first consider a system with one degree of freedom so that we only have to look for 1 variable $q(t)$.

If we assume that $q(t)$ is already the function that minimizes the action, this means that any deviation from it would no longer minimize S. To look into this, we consider a deviation $\delta q(t)$ that is small in the interval $t \in [t_1, t_2]$ such that it can be written as

$$q(t) + \delta q(t) \tag{2.29}$$

By virtue of us fixing the end points, we see that an immediate result of this is

$$\delta q(t_1) = \delta q(t_2) = 0 \tag{2.30}$$

As such, the difference in the action of the optimal and non-optimal path would be

$$\int_{t_1}^{t_2} \mathcal{L}(q + \delta q, \dot{q} + \delta \dot{q}, t)dt - \int_{t_1}^{t_2} \mathcal{L}(q, \dot{q}, t)dt \tag{2.31}$$

For convenience, we will label the non-optimal path as S' and let $S' - S = \delta S$ which is equal to the integral above. Since δq is taken to be a small deviation, we perform a Taylor expansion on the $\mathcal{L}$ in S' up to the first order term.

$$\mathcal{L}(q+\delta q, \dot{q}+\delta\dot{q}, t) \approx \mathcal{L}(q,\dot{q},t) + \Big(\frac{\partial \mathcal{L}}{\partial q}\delta q + \frac{\partial \mathcal{L}}{\partial \dot{q}}\delta\dot{q}\Big) + ... \tag{2.32}$$

$$\Rightarrow\ S' = \int_{t_1}^{t_2} \mathcal{L}(q+\delta q, \dot{q}+\delta\dot{q}, t)dt \tag{2.33}$$

$$\Rightarrow\ \delta S = S' - S \approx \int_{t_1}^{t_2} \Big(\frac{\partial \mathcal{L}}{\partial q}\delta q + \frac{\partial \mathcal{L}}{\partial \dot{q}}\delta\dot{q}\Big)dt \tag{2.34}$$

And by construction, if we find (2.34) such that it vanishes, we have found the path of least action. Hence, we now look to solve for $\delta S = 0$. First, we employ a trick from calculus which is a specific form of the *product rule*. In our context, it is written as such.

$$\frac{d}{dt}\Big(\frac{\partial \mathcal{L}}{\partial \dot{q}}\delta q\Big) = \delta q \frac{d}{dt}\Big(\frac{\partial \mathcal{L}}{\partial \dot{q}}\Big) + \frac{\partial \mathcal{L}}{\partial \dot{q}}\delta\dot{q} \tag{2.35}$$

$$\Rightarrow\ \delta S \approx \int_{t_1}^{t_2} dt\Big(\frac{\partial \mathcal{L}}{\partial q} - \frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{q}}\Big)\delta q + \int_{t_1}^{t_2} dt \frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{q}}\delta q \tag{2.36}$$

$$\Rightarrow\ \delta S \approx \int_{t_1}^{t_2} dt\Big(\frac{\partial \mathcal{L}}{\partial q} - \frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{q}}\Big)\delta q + \frac{\partial \mathcal{L}}{\partial \dot{q}}\delta q\Big|_{t_1}^{t_2} \tag{2.37}$$

But since $\delta q = 0$ when evaluated at the end points where $t = t_1, t_2$, the second term in δS vanishes.

$$\delta S = \int_{t_1}^{t_2} dt\Big(\frac{\partial \mathcal{L}}{\partial q} - \frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{q}}\Big)\delta q = 0 \tag{2.38}$$

$$\Rightarrow\ \frac{\partial \mathcal{L}}{\partial q} - \frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{q}} = 0 \tag{2.39}$$

Which brings us once again to the Euler-Lagrange equation (2.27). Here we assumed that the system we were working with had one degree of freedom but to extend this to a system with N degrees of freedom, we simply add the lower indices to our coordinates as done in (2.27).

Now, we move on to finding the functional form of the Lagrangian. Consider a system of free particles ($V(q_j) = 0$). We will look at a couple of properties of the Lagrangian to easier allow us to construct it.

Properties:

1. The Lagrangian is a **scalar** function. This follows from the nature of which it is defined.

2. The Lagrangian scaled by a constant factor α gives us the same equations of motions and thus the same physics. The constant will drop out in the Euler-Lagrange equation.

3. Adding a scalar constant to the Lagrangian itself gives us the same equations of motion. $\mathcal{L} \sim \mathcal{L}' = \mathcal{L} + \alpha$

4. if we define $\mathcal{L}' = \mathcal{L} + \frac{d}{dt}f(q,t)$, this does not change the physics either since

$$\begin{aligned} S' &= \int_{t_1}^{t_2} \mathcal{L}' dt = \int_{t_1}^{t_2} \big(\mathcal{L} + \frac{d}{dt}f(q,t)\big) dt \\ &= \int_{t_1}^{t_2} \mathcal{L} dt + f(q,t)|_{t_1}^{t_2} \\ &= S + \text{constant} \end{aligned}$$

and adding constants do not change the minimizing problem.

To better understand property 3, we shall look at a couple of examples of functions that satisfy this form.

Example

Here, we will provide 3 examples of functions that satisfy $g(q,\dot{q},t) = \frac{d}{dt}f(q,t)$. Note that these functions do not have to contain all the arguments specified in the given equation, but **cannot** have any additional arguments.

1. $g = g(t)$, (functions only dependent on time).
   $\Rightarrow f = \int h(t)dt = f(t)$, which is also just a function of time and satisfies the form.
2. $g = \alpha$, (constants).
   $\Rightarrow f = \int \alpha dt = \alpha t + C$, which is also just simply a function of time alone.
3. $g = h(q)\dot{q}$ where $h(q) = \frac{\partial f(q)}{\partial q}$, (an arbitrary function of q multiplied by $\dot{q}$).
   $\Rightarrow \frac{d}{dt}f(q) = \frac{\partial f(q)}{\partial t} + \frac{\partial f(q)}{\partial q}\dot{q} = 0 + \frac{\partial f(q)}{\partial q}\dot{q} = h(q)\dot{q}$

Continuing our construction of $\mathcal{L}$, we note that since we are dealing with a free particle. $\Rightarrow \mathcal{L}(q,\dot{q},t) \to \mathcal{L}(\dot{q},t)$ because space is *homogeneous* and physics does not depend on the position in which we perform it. We further assert that time is also homogeneous, $\Rightarrow \mathcal{L}(\dot{q},t) \to \mathcal{L}(\dot{q})$. With these 2 properties of the Lagrangian, (2.27) becomes

$$\frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{q}} = 0 \tag{2.40}$$

$$\Rightarrow \quad \frac{\partial \mathcal{L}}{\partial \dot{q}} = \text{constant} \tag{2.41}$$

$$\Rightarrow \quad \text{Either: } \begin{cases} 1) & \mathcal{L} \propto \dot{q} \\ 2) & \dot{q} \text{ is constant.} \end{cases} \tag{2.42}$$

But another property of space that we assert is that it is *isotropic* and physics does not depend on the direction in which perform it. $\Rightarrow \mathcal{L}(\dot{q}) \to \mathcal{L}(|\dot{q}|) \to \mathcal{L}(|\dot{q}|^2)$. This means that we **cannot** have a linear $\dot{q}$ term in our Lagrangian, $\Rightarrow \mathcal{L} \not\propto \dot{q}$. Hence we conclude that $\dot{q}$ is constant, which makes sense since it exactly agrees with Newton's 1$^{\text{st}}$ Law.

Knowing this, without loss of generality we can write $\mathcal{L}$ as

$$\mathcal{L} = a\dot{q}^2 + b\dot{q}^3 + c\dot{q}^4 + ... \tag{2.43}$$

and by virtue of the isotropy of space, this mandates that all odd powers of $\dot{q}$ be dropped. To attain the complete form, we have to look at *Galilean transformations* on the system. In fact for this part of the analysis, we can look at a system with an arbitrary potential.

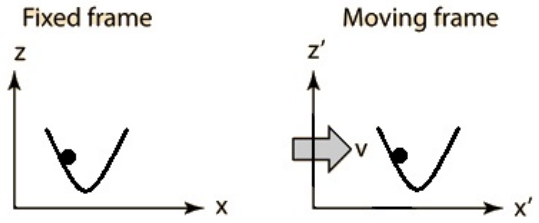

Figure 2.2: Galilean Boost of Mass in Potential Well System

As shown in figure 2.2, the boosted frame is denoted with primed coordinates. We will denote stationary and the boosted frame's potential as

$$V = V(\vec{r}, \dot{\vec{r}}) \tag{2.44}$$
$$V' = V(\vec{r}\,', \dot{\vec{r}}\,') \tag{2.45}$$

respectively. For generality, we will denote the general coordinate of the boost to be $\vec{r}$. Consider a boost by a small velocity $\vec{v}$. As such,

$$\vec{r}\,' = \vec{r} + \vec{v}t \tag{2.46}$$
$$\Rightarrow \quad \dot{\vec{r}}\,' = \dot{\vec{r}} + \vec{v} \tag{2.47}$$
$$\mathcal{L} = \mathcal{L}_{\text{free particle}}\left(\left|\dot{\vec{r}}\right|^2\right) + \alpha V(\vec{r}, \dot{\vec{r}}) \tag{2.48}$$
$$\mathcal{L}\,' = \mathcal{L}_{\text{free particle}}\left(\left|\dot{\vec{r}}\,'\right|^2\right) + \alpha V(\vec{r}\,', \dot{\vec{r}}\,') \tag{2.49}$$

Here we are using the result from our previous analysis asserting the isotropy of space, and we are assuming that the Lagrangian contains some addition of the potential.

$$\Rightarrow \quad \mathcal{L} - \mathcal{L}\,' = \mathcal{L}_{\text{FP}}\left(\left|\dot{\vec{r}}\right|^2\right) - \mathcal{L}_{\text{FP}}\left(\left|\dot{\vec{r}}\,'\right|^2\right) - \alpha\big(V(\vec{r}, \dot{\vec{r}}) - V(\vec{r}\,', \dot{\vec{r}}\,')\big)$$
$$\Rightarrow \quad \mathcal{L} - \mathcal{L}\,' = \mathcal{L}_{\text{FP}}\left(\left|\dot{\vec{r}}\right|^2\right) - \mathcal{L}_{\text{FP}}\left(\left|\dot{\vec{r}}\,'\right|^2\right)$$

This result follows from property 3 on Lagrangians. Since we are cosidering a small velocity boost, we can perform a Taylor expansion on the $\mathcal{L}\,'$ term to give

$$\begin{aligned}
\mathcal{L}_{\text{FP}}\left(\left|\dot{\vec{r}}\,'\right|^2\right) &= \mathcal{L}_{\text{FP}}\left(\dot{r}^2 + 2\dot{\vec{r}}\cdot\vec{v} + v^2\right) \\
&\approx \mathcal{L}_{\text{FP}}(\dot{r}^2) + \frac{\partial \mathcal{L}_{\text{FP}}(\dot{r}^2)}{\partial \dot{r}^2}(2\dot{\vec{r}}\cdot\vec{v} + v^2) \\
\Rightarrow \mathcal{L}\,' - \mathcal{L} &= \frac{\partial \mathcal{L}_{\text{FP}}(\dot{r}^2)}{\partial \dot{r}^2}(2\dot{\vec{r}}\cdot\vec{v} + v^2) \\
&= \frac{d}{dt}\left(\frac{\partial \mathcal{L}_{\text{FP}}(\dot{r}^2)}{\partial \dot{r}^2}(2\vec{r}\cdot\vec{v} + v^2)\right) - \left(\frac{d}{dt}\frac{\partial \mathcal{L}_{\text{FP}}(\dot{r}^2)}{\partial \dot{r}^2}\right)(2\vec{r}\cdot\vec{v} + v^2)
\end{aligned}$$

Now as a clarification on what we are working toward here, we define a function f as

$$f \equiv \frac{\partial \mathcal{L}_{\text{FP}}(\dot{r}^2)}{\partial \dot{r}^2}(2\vec{r}\cdot\vec{v}+v^2) \tag{2.50}$$

$$\Rightarrow \ \mathcal{L}\,' = \mathcal{L} + \frac{d}{dt}\big(f\big) - \Big(\frac{d}{dt}\frac{\partial \mathcal{L}_{\text{FP}}(\dot{r}^2)}{\partial \dot{r}^2}\Big)(2\vec{r}\cdot\vec{v}+v^2) \tag{2.51}$$

Then in order for us to get from (2.51) to the form in property 3 that will imply an equivalence of the boosted system, we need the last term on the right to vanish.

$$\Rightarrow \ \Big(\frac{d}{dt}\frac{\partial \mathcal{L}_{\text{FP}}(\dot{r}^2)}{\partial \dot{r}^2}\Big)(2\vec{r}\cdot\vec{v}+v^2) = 0$$

$$\Rightarrow \ \frac{\partial \mathcal{L}_{\text{FP}}(\dot{r}^2)}{\partial \dot{r}^2} = \text{constant}$$

$$\Rightarrow \ \mathcal{L}_{\text{FP}}(\dot{r}^2) \propto \dot{r}^2 \tag{2.52}$$

Which shows that the Lagrangian for an arbitrary system of 1 degree of freedom can be written as

$$\mathcal{L}(\dot{r}^2) = K\dot{r}^2 + \alpha V(\vec{r},\dot{\vec{r}}) \tag{2.53}$$

Then to find these constants, we just have to work out one problem and vary these unknown parameters until the correct equations of motions are reached which we have already done in section 2.2. Hence in generalized coordinates,

$$K = \frac{1}{2}m \tag{2.54}$$

$$\alpha = -1 \tag{2.55}$$

$$\boxed{\mathcal{L}(q,\dot{q}^2,t) = \frac{1}{2}m\dot{q}^2(t) - V\big(q(t)\big)} \tag{2.56}$$

Which is the Lagrangian derived in section 2.2.

Before we go deeper into Lagrangian mechanics, it is a good time to look at common coordinate transformations and calculus identities that we have and will be using as we go along.

Cylindrical Coordinates:

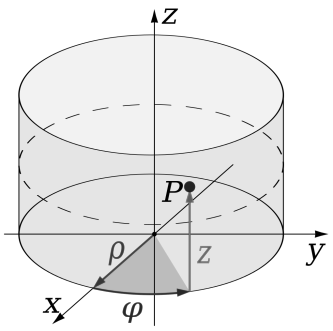

Figure 2.3: Cylindrical Coordinates

In physics, it is common consensus to take ϕ as the azimuthal angle around the equator. As such,

$$x = \rho\cos(\phi), \quad y = \rho\sin(\phi), \quad z = z \tag{2.57}$$

$$\rho = \sqrt{x^2 + y^2}, \quad \phi = \tan^{-1}(\frac{y}{x}) \tag{2.58}$$

$$\hat{\rho} = \cos(\phi)\hat{x} + \sin(\phi)\hat{y} \tag{2.59}$$

$$\vec{P} = \rho\hat{\rho} + z\hat{z} \tag{2.60}$$

$$\dot{\hat{\rho}} = -\sin(\phi)\dot{\phi}\hat{x} + \cos(\phi)\dot{\phi}\hat{y} \tag{2.61}$$

$$\dot{\vec{P}} = \dot{\rho}\hat{\rho} + \dot{z}\hat{z} \tag{2.62}$$

Spherical Coordinates:

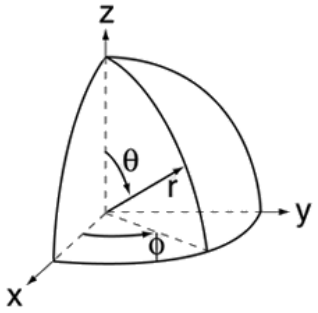

Figure 2.4: Spherical Coordinates

ϕ denotes the azimuthal angle around the equator and θ denotes the angle of inclination from the vertical.

$$x = r\sin(\theta)\cos(\phi), \quad y = r\sin(\theta)\sin(\phi), \quad z = r\cos(\theta) \tag{2.63}$$

$$r = \sqrt{x^2 + y^2 = z^2}, \quad \theta = \cos^{-1}(\frac{z}{r}), \quad \phi = \tan^{-1}(\frac{y}{x}) \tag{2.64}$$

$$\hat{r} = \sin(\theta)\big(\cos(\phi)\hat{x} + \sin(\phi)\hat{y}\big) + \cos(\theta)\hat{z} \tag{2.65}$$

$$\hat{\theta} = \cos(\theta)\big(\cos(\phi)\hat{x} + \sin(\phi)\hat{y}\big) - \sin(\theta)\hat{z} \tag{2.66}$$

$$\hat{\phi} = -\sin(\phi)\hat{x} + \cos(\phi)\hat{y} \tag{2.67}$$

$$\dot{\hat{r}} = \dot{\theta}\hat{\theta} + \sin(\theta)\dot{\phi}\hat{\phi} \tag{2.68}$$

$$\dot{\hat{\theta}} = \dot{\theta}\hat{r} + \cos(\theta)\dot{\phi}\hat{\phi} \tag{2.69}$$

$$\dot{\hat{\phi}} = -\big(\sin(\theta)\hat{r} + \cos(\theta)\hat{\theta}\big)\dot{\phi} \tag{2.70}$$

Center of Mass Coordinates:

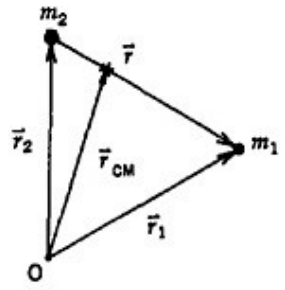

Figure 2.5: Center of Mass Coordinates

These coordinates are usually used for systems with a level of spherical symmetry, such as central potential and 2-body problems.

$$\mu = \frac{\prod_j m_j}{\sum_j m_j}, \quad M = \sum_j m_j \tag{2.71}$$

$$\vec{r}_{\text{CM}} = \frac{\sum_j m_j \vec{r}_j}{\sum_j m_j}, \quad \vec{v}_{\text{CM}} = \frac{\sum_j m_j \dot{\vec{r}}_j}{\sum_j m_j} \tag{2.72}$$

μ is known as the *reduced mass*, M is the *total mass* of the system and $\vec{r}_{\text{CM}}$ denotes the position of the system's center of mass from the original coordinate origin. With these coordinates, it makes things convenient as we treat the entirety of the system to be centered around its center of mass and all events are taken relative to that.

Useful Calculus Identities:

1. Total Derivative: $\frac{d}{dt}f(q,\dot{q},t) = \frac{\partial f}{\partial t} + \dot{q}\frac{\partial f}{\partial q} + \ddot{q}\frac{\partial f}{\partial \dot{q}}$
2. Chain Rule: $\frac{d}{dx}f\big(a(x),b(x)\big) = \frac{da(x)}{dx}\frac{\partial f}{\partial a} + \frac{db(x)}{dx}\frac{\partial f}{\partial b}$
3. Product Rule: $b(x)\frac{da(x)}{dx} = \frac{d}{dx}\big(a(x)b(x)\big) - a(x)\frac{db(x)}{dx}$
4. Integration by parts: $\int_{x_1}^{x_2} b(x)\big(\frac{da(x)}{dx}\big)dx = a(x)b(x)|_{x_1}^{x_2} - \int_{x_1}^{x_2} a(x)\big(\frac{db(x)}{dx}\big)dx$

Now that we have formulated the Lagrangian and the Euler-Lagrange equation, it would be good to define a systematic **work-flow** when approaching Lagrangian mechanics problems. The sequence of steps is as such.

1. Set-up the coordinate system.
2. Determine the Lagrangian $\mathcal{L}(q,\dot{q},t)$ of the system.
3. Compute the forces F_j and momenta p_j to formulate the equations of motions using the Euler-Lagrange equations (2.27).

4. Solve the equations of motion to see how the system evolves in time.

For a complicated dynamical system, step 4 is often done numerically since analytical solutions are often difficult to find.

When solving Lagrangian mechanics problems, our system is defined by a given potential which arises from the set-up. This is usually simple to formulate whereas the kinetic energy of the system is what is hard to derive. A general form of the kinetic energy can be written as such.

$$T = \frac{1}{2}\sum_j \sum_k a_{jk}\dot{q}_j\dot{q}_k, \quad a_{jk} = \sum_i m_i \frac{\partial r_i}{\partial q_j}\frac{\partial r_i}{\partial q_k} \tag{2.73}$$

It is best now to try several examples to make sure we understand how to use the new tools we are now equipped with.

Example 1

Consider a particle of mass m falling in a gravitational field with gravitational acceleration g. This problem is relatively trivial in Cartesian coordinates, so we will attempt this problem in polar coordinates (2-D spherical coordinates using r and ϕ). We then list down the necessary coordinates for our problem.

$$\vec{q} \equiv \{r, \phi\} \tag{2.74}$$

$$x = r\cos(\phi), \quad y = r\sin(\phi) \tag{2.75}$$

$$\Rightarrow \quad \frac{\partial x}{\partial r} = \cos(\phi), \quad \frac{\partial y}{\partial r} = \sin(\phi), \tag{2.76}$$

$$\frac{\partial x}{\partial \phi} = -r\sin(\phi), \quad \frac{\partial y}{\partial \phi} - r\cos(\phi) \tag{2.77}$$

Then using the formula for the generalized kinetic energy (2.73), we get

$$\begin{aligned} T &= \frac{1}{2}m\Big(\big(\frac{\partial x}{\partial r}^2 + \frac{\partial y}{\partial r}^2\big)\dot{r}^2 + 2\big(\frac{\partial x}{\partial r}\frac{\partial x}{\partial \phi} + \frac{\partial y}{\partial r}\frac{\partial y}{\partial \phi}\big)\dot{r}\dot{\phi} + \big(\frac{\partial x}{\partial \phi}^2 + \frac{\partial y}{\partial \phi}^2\big)\dot{\phi}^2\Big) \\ &= \frac{1}{2}m\Big(\big(\cos^2(\phi) + \sin^2(\phi)\big)(\dot{r}^2 + r^2\dot{\phi}^2) + 2\big(-r\sin(\phi)\cos(\phi) + r\sin(\phi)\cos(\phi)\big)\dot{r}\dot{\phi}\Big) \\ &= \frac{1}{2}m(\dot{r}^2 + r^2\dot{\phi}^2) \end{aligned} \tag{2.78}$$

After this, the next step would be to formulate the Lagrangian from T and V.

$$V = mgy = mgr\sin(\phi) \tag{2.79}$$

$$\Rightarrow \quad \mathcal{L} = T - V = \frac{1}{2}m(\dot{r}^2 + r^2\dot{\phi}^2) - mgr\sin(\phi) \tag{2.80}$$

From here, we compute the generalized forces and momentum.

$$F_r = \frac{\partial \mathcal{L}}{\partial r} = mr\dot{\phi}^2 - mg\sin(\phi) \tag{2.81}$$

$$p_r = \frac{\partial \mathcal{L}}{\partial \dot{r}} = m\dot{r} \tag{2.82}$$

$$F_\phi = \frac{\partial \mathcal{L}}{\partial \phi} = -mgr\cos(\phi) \tag{2.83}$$

$$p_\phi = \frac{\partial \mathcal{L}}{\partial \dot{\phi}} = mr^2\dot{\phi} \tag{2.84}$$

Plugging these into the Euler-Lagrange equations, we get

$$\begin{aligned} & F_r = \dot{p}_r \\ \Rightarrow\ & mr\dot{\phi}^2 - mg\sin(\phi) = m\ddot{r} \\ \Rightarrow\ & \boxed{\ddot{r} = r\dot{\phi}^2 - g\sin(\phi)} \end{aligned} \tag{2.85}$$

$$\begin{aligned} & F_\phi = p_\phi \\ \Rightarrow\ & -mgr\cos(\phi) = mr^2\ddot{\phi} + 2mr\dot{r}\dot{\phi} \\ \Rightarrow\ & \boxed{\ddot{\phi} = -\frac{g}{r}\cos(\phi) - \frac{2}{r}\dot{r}\dot{\phi}} \end{aligned} \tag{2.86}$$

Example 2

Consider a massless wheel of radius a that is rotating with angular frequency $\omega = \dot{\theta}$ with a massless rod of length l attached to the edge of it that is free to rotate about its attachment. There is a ball of mass m attached to the end of the rod. This is a 2 dimensional problem.

We take the centre of the wheel as the origin of our Cartesian coordinates (positive x is rightward, positive y is downward). We also define ϕ as the angle from the horizontal made by the rod with respect to its attachment. The position of the mass is thus given by

$$x = a\cos(\omega t) + l\sin(\phi) \tag{2.87}$$

$$y = -a\sin(\omega t) + l\cos(\phi) \tag{2.88}$$

$$\Rightarrow\ \dot{x} = -a\omega\sin(\omega t) + l\cos(\phi)\dot{\phi} \tag{2.89}$$

$$\Rightarrow\ \dot{y} = a\omega\cos(\omega t) - l\sin(\phi)\dot{\phi} \tag{2.90}$$

The terms in the Lagrangian are then,

$$\mathcal{L} = T - V = \frac{1}{2}m(\dot{x}^2 + \dot{y}^2) + mgy \tag{2.91}$$

Then looking at the individual terms,

$$\begin{aligned} T &= \frac{1}{2}m\Big(a^2\omega^2\sin^2(\omega t) - 2a\omega l\sin(\omega t)\cos(\phi)\dot{\phi} + l^2\cos^2(\phi)\dot{\phi}^2 \\ &\quad + \ a^2\omega^2\cos^2(\omega t) - 2a\omega l\cos(\omega t)\sin(\phi)\dot{\phi} + l^2\sin^2(\phi)\dot{\phi}^2\Big) \\ &= \frac{1}{2}m\big(a^2\omega^2 + 2a\omega l\sin(\phi - \omega t)\dot{\phi} + l^2\dot{\phi}\big) \end{aligned} \tag{2.92}$$

$$V = mg\left(a\sin(\omega t) - l\cos(\phi)\right) \tag{2.93}$$

From property 3 and 4 about Lagrangians, we can drop the constant and strictly time-dependent terms and still arrive at the same equations of motion. Hence, T and V reduce to

$$T = \frac{1}{2}m\big(2a\omega l\sin(\phi - \omega t)\dot{\phi} + l^2\dot{\phi}\big) \tag{2.94}$$

$$V = -mgl\cos(\phi) \tag{2.95}$$

We can further simplify T by using the identity

$$\begin{aligned}
&\frac{d}{dt}\cos(\phi-\omega t) = -\sin(\phi-\omega t)(\dot{\phi}-\omega) \\
\Rightarrow\ &\sin(\phi-\omega t)\dot{\phi} = \omega\sin(\phi-\omega t) - \frac{d}{dt}\cos(\phi-\omega t) \\
\Rightarrow\ &T = \frac{1}{2}m\big(2a\omega l\sin(\phi-\omega t) + l^2\dot{\phi}^2\big) \qquad (2.96)\\
\Rightarrow\ &\mathcal{L} = \frac{1}{2}m\big(2a\omega l\sin(\phi-\omega t) + l^2\dot{\phi}^2\big) + mgl\cos(\phi) \qquad (2.97)
\end{aligned}$$

Then to find the relevant equations of motion,

$$F_\phi = \frac{\partial\mathcal{L}}{\partial\phi} = ml\big(a\omega^2\cos(\phi-\omega t) - g\sin(\phi)\big) \qquad (2.98)$$

$$p_\phi = \frac{\partial\mathcal{L}}{\partial\dot{\phi}} = ml^2\dot{\phi} \qquad (2.99)$$

and from the Euler-Lagrange equation,

$$\begin{aligned}
&ml^2\ddot{\phi} = ml\big(a\omega^2\cos(\phi-\omega t) - g\sin(\phi)\big) \\
\Rightarrow\ &\boxed{\ddot{\phi} = \frac{1}{l}\big(a\omega^2\cos(\phi-\omega t) - g\sin(\phi)\big)} \qquad (2.100)
\end{aligned}$$

Example 3

Consider a hoop of mass m and radius R rolling down a ramp of mass M and inclination α that is free to move along the x-direction (horizontal axis). We take rightward as the positive x-direction and upward as the positive y-direction. Also let θ be the angle that denotes the angular position of the hoop. This is a 2 dimensional problem.

For clarity, we denote all coordinates to do with the hoop with a h subscript and coordinates to do with the ramp with no subscript. The coordinates of the hoop are then

$$\begin{aligned}
x_h &= R\theta\cos(\alpha) + x \qquad (2.101)\\
y_h &= R\theta\sin(\alpha) \qquad (2.102)\\
\Rightarrow\ \dot{x}_h &= R\dot{\theta}\cos(\alpha) + \dot{x} \qquad (2.103)\\
\Rightarrow\ \dot{y}_h &= R\dot{\theta}\sin(\alpha) \qquad (2.104)
\end{aligned}$$

Next we find the terms that constitute the Lagrangian.

$$T = \frac{1}{2}M\dot{x}^2 \tag{2.105}$$

$$\begin{aligned} T_h &= \frac{1}{2}m\big(\dot{x}_h^2 + \dot{y}_h^2\big) \\ &= \frac{1}{2}m\big(\dot{x}^2 + 2R\dot{\theta}\dot{x}\cos(\alpha) + R^2\dot{\theta}^2\big) \end{aligned} \tag{2.106}$$

$$V = 0 \tag{2.107}$$

$$V_h = mgy_h = -mgR\theta\sin(\alpha) \tag{2.108}$$

$$\begin{aligned} \Rightarrow\ \mathcal{L} &= (T + T_h) - (V + V_h) \\ &= \frac{1}{2}M\dot{x}^2 + \frac{1}{2}m\big(\dot{x}^2 + 2R\dot{\theta}\dot{x}\cos(\alpha) + R^2\dot{\theta}^2\big) + mgR\theta\sin(\alpha) \end{aligned} \tag{2.109}$$

With this, we find the generalized forces and momenta.

$$\frac{\partial\mathcal{L}}{\partial x} = 0, \quad \frac{\partial\mathcal{L}}{\partial\theta} = mgR\sin(\alpha) \tag{2.110}$$

$$\frac{\partial\mathcal{L}}{\partial\dot{x}} = M\dot{x} + m\dot{x} + mR\dot{\theta}\cos(\alpha), \quad \frac{\partial\mathcal{L}}{\partial\dot{\theta}} = mR\dot{x}\cos(\alpha) + mR^2\dot{\theta} \tag{2.111}$$

$$\Rightarrow\ (M+m)\ddot{x} + mR\ddot{\theta}\cos(\alpha) = 0$$

$$\Rightarrow\ \boxed{\ddot{x} = -\ddot{\theta}\Big(\frac{mR}{M+m}\Big)\cos(\theta)} \tag{2.112}$$

$$\begin{aligned} &mR\ddot{x}\cos(\alpha) + mR^2\ddot{\theta} = mgR\sin(\alpha) \\ \Rightarrow\ &\ddot{\theta}\Big(R^2 - \big(\frac{m}{M+m}\big)R^2\cos^2(\alpha)\Big) = gR\sin(\alpha) \\ \Rightarrow\ &\boxed{\ddot{\theta} = \frac{(M+m)g\sin(\alpha)}{MR + mR\sin^2(\alpha)} = \text{constant}} \end{aligned} \tag{2.113}$$

Since it is easy to solve for a variable that is undergoing constant acceleration, we will continue to provide the solution for $\theta(t)$ for this problem.

$$\theta(t) = \left(\frac{(M+m)g\sin(\alpha)}{MR + mR\sin^2(\alpha)}\right)t^2 + \omega_0 t \tag{2.114}$$

Example 4

Consider a simple pendulum system of mass m, but instead of a string, we replace it with a spring with spring constant k and rest length l. θ defines the deviation of the spring from the vertical axis and r denotes the distance of the mass away from the origin.

Setting up the necessary coordinates for the problem, we define the origin as the attachment point of the spring. Rightward is defined as the positive x-direction and downward

as the positive y-direction.

$$x = r\sin(\theta) \tag{2.115}$$
$$y = r\cos(\theta) \tag{2.116}$$
$$\Rightarrow\ \dot{x} = \dot{r}\sin(\theta) + r\cos(\theta)\dot{\theta} \tag{2.117}$$
$$\Rightarrow\ \dot{y} = \dot{r}\cos(\theta) - r\sin(\theta)\dot{\theta} \tag{2.118}$$

Formulating the Lagrangian,

$$T = \frac{1}{2}m(\dot{x}^2 + \dot{y}^2) = \frac{1}{2}m(\dot{r}^2 + r^2\dot{\theta}^2) \tag{2.119}$$
$$V = mgy + \frac{1}{2}k(l-r)^2 = -mgr\cos(\theta) + \frac{1}{2}k(l-r)^2 \tag{2.120}$$
$$\Rightarrow\ \mathcal{L} = \frac{1}{2}m(\dot{r}^2 + r^2\dot{\theta}^2) + mgr\cos(\theta) - \frac{1}{2}k(l-r)^2 \tag{2.121}$$

The generalized forces and momenta are,

$$\frac{\partial\mathcal{L}}{\partial r} = mr\dot{\theta}^2 + mg\cos(\theta) + k(l-r), \quad \frac{\partial\mathcal{L}}{\partial\dot{r}} = m\dot{r} \tag{2.122}$$
$$\frac{\partial\mathcal{L}}{\partial\theta} = -mgr\sin(\theta), \quad \frac{\partial\mathcal{L}}{\partial\dot{\theta}} = mr^2\dot{\theta} \tag{2.123}$$
$$\Rightarrow\ \boxed{\ddot{r} = r\dot{\theta}^2 + g\cos(\theta) + \frac{k}{m}(l-r)} \tag{2.124}$$
$$\Rightarrow\ 2mr\dot{r}\dot{\theta} + mr^2\ddot{\theta} = -mgr\sin(\theta)$$
$$\Rightarrow\ \boxed{\ddot{\theta} = -\frac{1}{r}\big(g\sin(\theta) + 2\dot{r}\dot{\theta}\big)} \tag{2.125}$$

Example 5

In this next example, we prove that the shortest path between 2 points in flat space is a straight line. We adopt Cartesian coordinates for this problem and begin by considering 2 arbitrary fixed points in 3-D space, $\vec{x}_1$ and $\vec{x}_2$. Then for any arbitrary path between these 2 points, an infinitesimal line segment along this path is given as

$$\begin{aligned} ds &= \sqrt{dx^2 + dy^2 + dz^2} \\ &= \sqrt{1 + \big(\frac{dy}{dx}\big)^2 + \big(\frac{dz}{dx}\big)^2} \end{aligned} \tag{2.126}$$

This would be the Lagrangian of our problem. As such, to apply Euler-Lagrange equations we look at its constituent terms.

$$\frac{\partial\mathcal{L}}{\partial y} = 0, \quad \frac{\partial\mathcal{L}}{\partial y'(x)} = y'(1 + y'^2 + z'^2)^{-\frac{1}{2}} \tag{2.127}$$
$$\frac{\partial\mathcal{L}}{\partial z} = 0, \quad \frac{\partial\mathcal{L}}{\partial z'(x)} = z'(1 + y'^2 + z'^2)^{-\frac{1}{2}} \tag{2.128}$$

We now look at the Euler-Lagrange equation for y, which gives

$$\frac{\partial \mathcal{L}}{\partial y} = \frac{d}{dx}\frac{\partial \mathcal{L}}{\partial y'} \tag{2.129}$$

$$\Rightarrow \frac{d}{dx}\frac{y'}{\sqrt{1+y'^2+z'^2}} = 0$$

$$\Rightarrow \frac{y'}{\sqrt{1+y'^2+z'^2}} = \text{constant} = \sqrt{C} \tag{2.130}$$

$$\Rightarrow y'^2 = C(1+y'^2+z'^2)$$

$$\Rightarrow y'^2 = \frac{C(1+z'^2)}{1-C} \equiv A + Bz'^2$$

$$\Rightarrow z'^2 = C + Dy'^2$$

$$\Rightarrow y'^2 = A + B(C+Dy'^2) \equiv \alpha + \beta y'^2$$

$$\Rightarrow \frac{dy}{dx} \text{ and } \frac{dz}{dx} = \text{constant}$$

As such we conclusively find that the principle of least action mandates the path taken between 2 points without constraint must be a straight line.

Chapter 3

Conserved Quantities

For a system with N degrees of freedom, there exists $2N-1$ independent functions of q_j and $\dot{q}_j$ that are constant throughout time despite its arguments being time varying. These are known as *integrals of motion*. Among these $2N-1$ integrals of motion, only some bear physical significance being derived from symmetries of space and time. We will explore these conserved quantities in this chapter.

§3.1 Conservation of Energy

The first conservation law we will look at arises from the homogeneity of time. This means that arbitrary translations in time do not change the physics of a dynamical system. We will look at the Newtonian and Lagrangian approach to revealing this integral of motion, as will be done for the other conserved quantities as well.

Newtonian Approach

Consider a particle of mass m that travels along a path between 2 points $\vec{x}_1$ and $\vec{x}_2$. Its infinitesimal line segment is denoted as $d\vec{r}$. The work done on the particle along this path is then

$$\begin{aligned} W_{1,2} &= \int_{x_1}^{x_2} \vec{F} \cdot d\vec{r} \qquad (3.1) \\ &= m \int_{x_1}^{x_2} \frac{d\vec{v}}{dt} \cdot \frac{d\vec{r}}{dt} dt \\ &= \frac{1}{2} m \int_{x_1}^{x_2} \frac{d}{dt} (\vec{v} \cdot \vec{v}) dt \\ &= \frac{1}{2} m \big(v^2(t_2) - v^2(t_1) \big) = \triangle T \qquad (3.2) \end{aligned}$$

We also assert that this is a non-dissipative system where we have a conservative potential.

As such,

$$W_{1,2} = \int_{x_1}^{x_2} \vec{F} \cdot d\vec{r} \tag{3.3}$$

$$= \int_{x_1}^{x_2} (- \triangle V) \cdot d\vec{r}$$

$$= V(x_1) - V(x_2) = - \triangle V \tag{3.4}$$

$$\Rightarrow \triangle T = - \triangle V$$

$$\Rightarrow \triangle (T + V) = \triangle E = 0 \tag{3.5}$$

From this, we see that $E = T + V$ is a conserved quantity which does not vary with time. Hence, the **total energy of the system is conserved**.

Lagrangian Approach

Here, be begin by asserting that time is homogeneous $\Rightarrow \frac{\partial \mathcal{L}}{\partial t} = 0$. We also first work with a system that has one degree of freedom. As such,

$$\frac{d}{dt}\mathcal{L}(q, \dot{q}, t) = \frac{\partial \mathcal{L}}{\partial t} + \dot{q}\frac{\partial \mathcal{L}}{\partial q} + \ddot{q}\frac{\partial \mathcal{L}}{\partial \dot{q}} \tag{3.6}$$

$$= 0 + \dot{q}\frac{\partial \mathcal{L}}{\partial q} + \ddot{q}\frac{\partial \mathcal{L}}{\partial \dot{q}}$$

$$= \dot{q}\Big(\frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{q}}\Big) + \Big(\frac{d}{dt}\dot{q}\Big)\frac{\partial \mathcal{L}}{\partial \dot{q}}$$

$$= \frac{d}{dt}\Big(\dot{q}\frac{\partial \mathcal{L}}{\partial \dot{q}}\Big) \tag{3.7}$$

In the above derivation, we utilized the Euler-Lagrange equation and the definition of the product rule. The result is that we have an unknown conserved quantity.

$$\frac{d}{dt}\Big(\dot{q}\frac{\partial \mathcal{L}}{\partial \dot{q}} - \mathcal{L}(q, \dot{q}, t)\Big) = 0 \tag{3.8}$$

To see what this quantity actually is, we extend it to a system of N degrees of freedom.

$$\Rightarrow \quad \frac{d}{dt}\Big(\sum_j^N \dot{q}_j \frac{\partial \mathcal{L}}{\partial \dot{q}_j} - \mathcal{L}(\vec{q}, \dot{\vec{q}}, t)\Big) = 0 \tag{3.9}$$

$$\Rightarrow \quad \sum_j^N \dot{q}_j \frac{\partial \mathcal{L}}{\partial \dot{q}_j} - \mathcal{L}(\vec{q}, \dot{\vec{q}}, t) = \text{constant} \tag{3.10}$$

Further investigating the first term, since

$$\mathcal{L} = T - V = \frac{1}{2}m\Big(\sum_{jk} a_{jk}\dot{q}_j\dot{q}_k\Big) - V(q) \tag{3.11}$$

$$\Rightarrow \quad \sum_j^N \dot{q}_j \frac{\partial \mathcal{L}}{\partial \dot{q}_j} = \sum_j \frac{1}{2}m\dot{q}_j \sum_{kl} a_{kl}(\delta_{jk}\dot{q}_l + \delta_{jl}\dot{q}_k) \tag{3.12}$$

$$= \frac{1}{2}m\sum_{jk} a_{jk}(\dot{q}_j\dot{q}_k + \dot{q}_k\dot{q}_j) = 2T \tag{3.13}$$

We have assumed here that V is strictly a function of q. Putting this back into our original integral of motion,

$$\Rightarrow \quad \sum_j^N \dot{q}_j \frac{\partial \mathcal{L}}{\partial \dot{q}_j} - \mathcal{L}(\vec{q}, \dot{\vec{q}}, t) = 2T - (T - V) = T + V$$

$$\Rightarrow \quad \boxed{E = \sum_j^N \dot{q}_j \frac{\partial \mathcal{L}}{\partial \dot{q}_j} - \mathcal{L}(\vec{q}, \dot{\vec{q}}, t)} \tag{3.14}$$

we find that the integral of motion we found was in fact the energy!

(3.14) gives the a *generalized form of the energy* of a system in the language of Lagrangian mechanics and, as proven, is a **conserved quantity**. Note that if we did not assume $V = V(q)$, our generalized energy would not have equated to just $T + V$, but contain an additional term.

$$\boxed{E = T + V - \sum_j \frac{\partial V(q, \dot{q})}{\partial \dot{q}_j}} \tag{3.15}$$

§3.2 Conservation of Linear Momentum

We will now look to derive the conservation law that arises from the homogeneity of space. This implies that arbitrary translations in space do not change the physics of a dynamical system. As before, we first look at the Newtonian approach to this result.

Newtonian Approach

By Newton's second law of motion,

$$\vec{F} = \dot{\vec{p}} \tag{3.16}$$

$$\Rightarrow \quad \triangle \vec{p} = \int_{t_1}^{t_2} \vec{F} dt \tag{3.17}$$

Following immediately from this, if

$$\vec{F} = \vec{0} \quad \Rightarrow \quad \triangle \vec{p} = \vec{0} \tag{3.18}$$

Which implies the **conservation of linear momentum**.

Lagrangian Approach

For the Lagrangian approach, proving it for a system with one degree of freedom is equally straight forward. Starting with the Euler-Lagrangian equation (2.27), we enforce that $F = 0$.

$$\Rightarrow \quad \frac{\partial \mathcal{L}}{\partial q} = 0 \tag{3.19}$$

$$\Rightarrow \quad \frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \dot{q}} = 0 \tag{3.20}$$

$$\Rightarrow \quad \frac{\partial \mathcal{L}}{\partial \dot{q}} = p = \text{constant} \tag{3.21}$$

Hence, linear momentum is conserved.

We now show that this holds for a system of N particles. We denote the coordinates of each particle as $\vec{r}$. Assuming that no net external force acts on the system,

$$\sum_j \vec{F}_j = \vec{0} \tag{3.22}$$

$$\Rightarrow \sum_j \nabla_j V(\vec{r}_1, ..., \vec{r}_N) = \sum_j \frac{\partial V}{\partial \vec{r}_j} = \vec{0} \tag{3.23}$$

Now, exploiting the homogeneity of space, the physics will not change if we translate the system by a differential amount ($\vec{r}\,' = \vec{r} + \vec{\epsilon}$).

$$\Rightarrow \mathcal{L} = \mathcal{L}' = \mathcal{L}(\vec{r} + \vec{\epsilon}, \dot{\vec{r}} + \dot{\vec{\epsilon}}, t) \tag{3.24}$$

$$= \mathcal{L}(\vec{r}, \dot{\vec{r}}, t) + \sum_j \Big(\frac{\partial \mathcal{L}}{\partial \vec{r}_j} \cdot \vec{\epsilon} + \frac{\partial \mathcal{L}}{\partial \dot{\vec{r}}_j} \cdot \dot{\vec{\epsilon}} \Big) + ... \tag{3.25}$$

Since $\vec{\epsilon}$ is constant $\Rightarrow \dot{\vec{\epsilon}} = 0$. Then taking from the Euler-Lagrange equation,

$$\mathcal{L}' \approx \mathcal{L}\ (\vec{r}, \dot{\vec{r}}, t) + \frac{d}{dt} \sum_j \Big(\frac{\partial \mathcal{L}}{\partial \dot{\vec{r}}_j} \cdot \vec{\epsilon} \Big) \tag{3.26}$$

$$\Rightarrow \mathcal{L}' - \mathcal{L} = \frac{d}{dt} \sum_j \Big(\frac{\partial \mathcal{L}}{\partial \dot{\vec{r}}_j} \cdot \vec{\epsilon} \Big)$$

$$= \frac{d}{dt} \sum_j (\vec{p}_j \cdot \vec{\epsilon}) = 0$$

$$\Rightarrow \frac{d}{dt} \Big(\sum_j \vec{p}_j \Big) = 0 \tag{3.27}$$

$$\Rightarrow \sum_j \vec{p}_j = \text{constant} \tag{3.28}$$

Hence we show that the **total linear momentum of the system is conserved**.

§3.3 Conservation of Angular Momentum

Another profoundly significant conservation law is one that arises from the isotropy of space. Isotropy implies that arbitrary rotations in space do not change the physics of a dynamical system (i.e. directionality is arbitrary).

Newtonian Approach

Starting from the notion that space is isotropic, we can arbitrarily pick a z-direction and consider only rotations on the x,y-plane without loss of generality. Then starting with an arbitrary vector in the x,y-plane, we rotate our system by a small angle ϕ, which causes our new coordinates (denoted with primes) to be

$$x' = x\cos(\phi) - y\sin(\phi) \tag{3.29}$$

$$y' = y\cos(\phi) + x\sin(\phi) \tag{3.30}$$

$$z' = z \tag{3.31}$$

We adopt the small angle approximation on ϕ with $\phi \ll 1$,

$$\Rightarrow \; x' = x - y\phi \tag{3.32}$$
$$\Rightarrow \; y' = y + x\phi \tag{3.33}$$

We can then write these small deviations from the original vector as

$$\vec{\epsilon} = \begin{bmatrix} -y\phi \\ x\phi \\ 0 \end{bmatrix} \tag{3.34}$$

noting that $\vec{\epsilon}$ is not a constant in this context. We then perform the Taylor expansion of the new rotated system's Lagrangian.

$$\begin{aligned} \mathcal{L}' &= \mathcal{L}(\vec{r} + \vec{\epsilon}, \dot{\vec{r}} + \dot{\vec{\epsilon}}) \\ &\approx \mathcal{L} - \frac{\partial \mathcal{L}}{\partial x} y\phi - \frac{\partial \mathcal{L}}{\partial \dot{x}} \dot{y}\phi + \frac{\partial \mathcal{L}}{\partial y} x\phi + \frac{\partial \mathcal{L}}{\partial \dot{y}} \dot{x}\phi \end{aligned}$$

Since we are asserting rotational invariance,

$$\Rightarrow \; \mathcal{L} = \mathcal{L}' \tag{3.35}$$
$$\Rightarrow \; -\frac{\partial \mathcal{L}}{\partial \dot{x}} \dot{y}\phi + \frac{\partial \mathcal{L}}{\partial y} x\phi + \frac{\partial \mathcal{L}}{\partial \dot{y}} \dot{x}\phi = 0$$
$$\Rightarrow \; (\dot{p}_y x + p_y \dot{x} - \dot{p}_x y - p_x \dot{y} = 0)\phi = 0$$
$$\Rightarrow \; \frac{d}{dt}(p_y x - p_x y) = \frac{d}{dt} L_z = 0 \tag{3.36}$$

Thus showing that angular momentum about the z-axis is conserved. Then extending this using again that space is isotropic, **angular momentum is conserved** about any axis. Beware about the similar notation of angular momentum and Lagrangians. Even though both are denoted with an L, different font styles are used to distinguish them and angular momentum is usually denoted as a vector or with a coordinate subscript.

These conservation laws as we have seen, have come about due to specific symmetries in our systems. These are not coincidences and was proven as such by Emmy Noether in 1915. The formal statement of *Neother's theorem* is given below.

Theorem 3.3.1. *Every differentiable symmetry of the action of a physical system has a corresponding conservation law.*

Proof. The proof that is presented here is not constructive. Consider a system with a set of coordinates $\{q_j\}$. We also have a parameter s such that $q_j = q_j(s)$. Additionally we define the quantity

$$H_j \equiv p_j \frac{\partial q_j}{\partial s} \tag{3.37}$$

Now consider the total time derivative of H_j.

$$\begin{aligned} \frac{d}{dt}H_j &= \dot{p_j}\frac{\partial q_j}{\partial s} + p_j\frac{d}{dt}\frac{\partial q_j}{\partial s} \qquad (3.38)\\ &= \frac{\partial \mathcal{L}}{\partial q_j}\frac{\partial q_j}{\partial s} + \frac{\partial \mathcal{L}}{\partial \dot{q_j}}\frac{\partial \dot{q_j}}{\partial s} \\ \Rightarrow \quad \frac{dH}{dt} &= \sum_j \frac{d}{dt}H_j = \sum_j \frac{\partial \mathcal{L}}{\partial q_j}\frac{\partial q_j}{\partial s} + \frac{\partial \mathcal{L}}{\partial \dot{q_j}}\frac{\partial \dot{q_j}}{\partial s} = \frac{d\mathcal{L}}{ds} \qquad (3.39) \end{aligned}$$

From this relation, we see that if s is a symmetry of the system, then $\frac{\partial \mathcal{L}}{\partial s} = 0$ which means H is now a conserved quantity of this system. Hence for every symmetry s, we get a conserved quantity H as Noether's theorem states. □

This theorem manifests itself time and time again in all areas of physics, and is an extremely fundamental theorem.

§3.4 Mechanical Similarity

We have seen before that an arbitrary scaling of the Lagrangian does not alter the resultant equations of motion. What about a scaling of the coordinates? Would this change the equations of motion? It turns out that if we have a potential that is a homogeneous function of the coordinates, then a specific scaling of the coordinates would ensure *geometrically similar* trajectories of the particles in the system.

Definition 3.4.1. Homogeneous Functions of Degree κ: *Given a function $f : X \to \mathbb{R}$, it is homogeneous of degree κ if $\forall x \in X$ and $\lambda > 0$ with $\lambda x \in X$,*

$$f(\lambda x) = \lambda^\kappa f(x) \qquad (3.40)$$

Definition 3.4.2. Geometrically Similar Paths: *Paths are said to be geometrically similar if their shape, defined by their outline or external boundary, are the same and differ only by uniform scaling.*

Also, a potential that is a homogeneous function of the system coordinates satisfies the following property.

$$V(\alpha\vec{r}_1, ..., \alpha\vec{r}_N) = \alpha^\kappa V(\vec{r}_1, ..., \vec{r}_N) \qquad (3.41)$$

Where α is an arbitrary constant and κ is called the *degree of homogeneity* of the function.

An example of this would be the simple harmonic oscillator potential, where $V(\vec{r}_1, \vec{r}_2) = \frac{1}{2}k|\vec{r}_1 - \vec{r}_2|^2$. Here, $\kappa = 2$ because the potential is a squared function of coordinates. For simplicity, we now look at a one degree of freedom system to do further analysis on this. Consider

the following coordinate transformation.

$$q \to q' = \alpha q \tag{3.42}$$
$$t \to t' = \beta t \tag{3.43}$$

Then taking the derivative of q' with respect to t' would yield,

$$\dot{q}' = \frac{d}{dt'}q' = \Big(\frac{\alpha}{\beta}\Big)\dot{q} \tag{3.44}$$
$$\begin{aligned} \Rightarrow \ \mathcal{L}' &= \mathcal{L}(q', \dot{q}') = \mathcal{L}(\alpha q, \Big(\frac{\alpha}{\beta}\Big)\dot{q}) \\ &= \frac{1}{2}m\dot{q}'^2 - V(q') \\ &= \frac{1}{2}m\Big(\frac{\alpha}{\beta}\dot{q}\Big)^2 - \alpha^\kappa V(q) \end{aligned} \tag{3.45}$$

However since we are looking for a global scaling of the Lagrangian, this means that we are looking for $\frac{\alpha}{\beta}$ such that $\mathcal{L}' = \alpha^\kappa \mathcal{L}$.

$$\begin{aligned} &\Rightarrow \ \Big(\frac{\alpha}{\beta}\Big)^2 = \alpha^\kappa \\ &\Rightarrow \ \boxed{\beta = \alpha^{1-\frac{\kappa}{2}}} \end{aligned} \tag{3.46}$$

Hence we get a global scaling of the Lagrangian which indeed does not alter the resultant equations of motion. Then given that we have a homogeneous potential of degree κ in the coordinates, the result of of scaling all the coordinates by the same factor α would yield geometrically similar paths as solutions to the equations of motion. The ratio of the scaled (primed) and unscaled (unprimed) times is thus given by,

$$\boxed{\frac{t'}{t} = \Big(\frac{l'}{l}\Big)^{1-\frac{\kappa}{2}}} \tag{3.47}$$

where l'/l is the ratio of dimensions of the scaled and unscaled paths. In fact **all** mechanical quantities at specified points in time would experience a scaling ratio equal to the ratio of l'/l raised to a function of κ. This is the principle of *mechanical similarity*. The associated relation of several mechanical quantities are given below.

$$\frac{v'}{v} = \Big(\frac{l'}{l}\Big)^{\frac{\kappa}{2}} \tag{3.48}$$
$$\frac{E'}{E} = \Big(\frac{l'}{l}\Big)^{\kappa} \tag{3.49}$$
$$\frac{L'}{L} = \Big(\frac{l'}{l}\Big)^{1+\frac{\kappa}{2}} \tag{3.50}$$

where v is velocity, E is energy and L is angular momentum.

We will now work through 2 examples to illustrate this in explicit systems.

Example 1

Consider again a harmonic oscillator quadratic potential written as

$$V(x) = \frac{1}{2}m\omega^2 x^2 \tag{3.51}$$

We also define the period of oscillation, amplitude and degree of the potential as

$$\text{Period} = \tau = \frac{2\pi}{\omega} \tag{3.52}$$
$$\text{Amplitude} = A \tag{3.53}$$
$$\text{Degree of Potential} = \kappa = 2 \tag{3.54}$$

Then if we scale the amplitude of oscillation by a factor α and the time by a factor β,

$$\Rightarrow \quad A' = \alpha A \tag{3.55}$$
$$\Rightarrow \quad \tau' = \beta\tau \tag{3.56}$$

But by the relation in (3.46) of mechanical similarity,

$$\beta = \alpha^{1-\frac{\kappa}{2}} = \alpha^{1-\frac{2}{2}} = \alpha^0 \tag{3.57}$$
$$\Rightarrow \quad \tau' = \tau \tag{3.58}$$

Hence we see that the period of oscillation of a simple harmonic oscillator does **not** depend on the amplitude of oscillation.

Example 2

In this next example, we point toward the discovery of *dark matter* using the concept of mechanical similarity. For this, we will be looking at Newtonian gravitational potentials and orbits of celestial bodies which we will treat as particles. But first, we look at a general situation of a power law potential

$$V(q) = aq^\kappa \tag{3.59}$$

where a is some arbitrary constant. Before we move on to diving into any deeper analysis, we first derive several useful mathematical identities.

$$q\frac{\partial V}{\partial q} = qa\kappa q^{\kappa-1} = \kappa V(q) \tag{3.60}$$
$$\frac{d}{dt}(pq) = \dot{p}q + \dot{q}p \tag{3.61}$$
$$T = \frac{1}{2}mv^2 = \frac{1}{2}p\dot{q} \tag{3.62}$$

Equation (3.60) is a result known as *Euler's theorem for homogeneous functions*. From these, we see that

$$\frac{d}{dt}(pq) = -q\frac{\partial V}{\partial q} + 2T = -\kappa V(q) + 2T \tag{3.63}$$

We now look at what is known as the *time-average* of a quantity. For a given quantity $f(t)$, its time-average is defined as

$$\langle f(t)\rangle = \lim_{\tau\to\infty}\frac{1}{\tau}\int_0^\tau f(t)dt \tag{3.64}$$

Hence if we look at the time average of $\frac{d}{dt}(pq)$, we get

$$\langle\frac{d}{dt}(pq)\rangle = \lim_{\tau\to\infty}\frac{1}{\tau}\int_0^\tau \frac{d}{dt}(pq)dt \tag{3.65}$$

$$= \lim_{\tau\to\infty}\frac{1}{\tau}(pq)|_0^\tau \tag{3.66}$$

Since we are looking at orbiting 'particles', this means that p and q remain finite at all times. which means the integral above evaluates to a finite value making $\langle\frac{d}{dt}(pq)\rangle = 0$.

$$\boxed{2\langle T\rangle = \kappa\langle V\rangle} \tag{3.67}$$

Now if we look at gravitational potential energy, it is given as

$$V(r) = -\frac{GMm}{r} \tag{3.68}$$

$$\Rightarrow\ V \propto \frac{1}{r}$$

$$\Rightarrow\ \kappa = -1$$

$$\Rightarrow\ \boxed{2\langle T\rangle = -\langle V\rangle} \tag{3.69}$$

This result in (3.69) is known as the *Virial theorem*. There is actually an even more generalized form of this theorem which we will derive later on.

Back to dark matter, what gives rise to its existence is the fact that when cosmologists have computed the kinetic energy of galaxies, the result is strangely that $2\langle T\rangle > -\langle V\rangle$! This means that there has to be a lot more mass than what is observed in the galaxy to sustain it moving with that much kinetic energy. This additional required matter is what we have found to be dark matter.

§3.5 Motion in 1-D

Now, we look at systems that may not exist in 1 dimensional space, but have only one effective degree of freedom such that we only require to specify one coordinate to describe the system. We start but writing the general form of energy for such a system.

$$E = \frac{1}{2}ma(q)\dot{q}^2 + V(q) \tag{3.70}$$

where $a(q) > 0$ is some function of the coordinate.

A specific example of this would be Cartesian coordinates, making $q = x$ and $a(q) = a(x) = 1$.

We can then manipulate the above equation as such.

$$\begin{aligned} E - V(q) &= \frac{1}{2} m a(q) \dot{q}^2 \\ \Rightarrow \quad \dot{q}^2 &= \frac{2\big(E - V(q)\big)}{m a(q)} > 0 \end{aligned} \tag{3.71}$$

This implies that a particle cannot exist in a region where $V(q) > E$. Hence for arbitrary potentials, if a particle exists in a local potential well without sufficient energy to escape it, it will remain *bound* within that potential well. The converse is known as *unbound*. Looking at a bound particle, if it is oscillating within the well, then the boundary points where the particle changes in its direction of motion are called *turning points*. Also if a particle happens to be exactly at a local maxima, then any slight perturbation would cause it to fall off in either direction. This is known as *spontaneous symmetry breaking*.

We will now work through the *2-body problem* and show how we can reduce it to be effectively a 1-D problem by utilizing the center of mass frame.

Example

We have previously touched on center of mass coordinates in section 2.4 but here we will show specific application of it and its derivation. You can refer to the diagram in section 2.4 to have a better visualization of the problem. The Lagrangian can immediately be written as

$$\mathcal{L} = \frac{1}{2} m_1 |\vec{r}|_1^2 + \frac{1}{2} m_2 |\vec{r}|_2^2 - V(\vec{r}_1 - \vec{r}_2) \tag{3.72}$$

The center of mass coordinate is then

$$\vec{r}_{CM} = \frac{m_1 \vec{r}_1 + m_2 \vec{r}_2}{M}, \quad M \equiv m_1 + m_2 \tag{3.73}$$

For convenience, wee also define another coordinate

$$\vec{r} \equiv \vec{r}_1 - \vec{r}_2 \tag{3.74}$$

$$\vec{r}_{1,CM} \equiv \vec{r}_1 - \vec{r}_{CM} = \Big(\frac{m_2}{M}\Big)\vec{r} \tag{3.75}$$

$$\vec{r}_{2,CM} \equiv \vec{r}_2 - \vec{r}_{CM} = \Big(\frac{m_1}{M}\Big)\vec{r} \tag{3.76}$$

$$\tag{3.77}$$

Plugging this into our terms in the Lagrangian, we get

$$\begin{aligned} T &= \frac{1}{2} m_1 \Big| \Big(\frac{m_2}{M}\Big)\dot{\vec{r}} + \dot{\vec{r}}_{CM} \Big|^2 + \frac{1}{2} m_2 \Big| \Big(\frac{m_1}{M}\Big)\dot{\vec{r}} + \dot{\vec{r}}_{CM} \Big|^2 \\ &= \frac{1}{2} m_1 \Big(\big(\frac{m_2}{M}\big)^2 \Big|\dot{\vec{r}}\Big|^2 + \Big|\dot{\vec{r}}_{CM}\Big|^2 + \frac{2m_2}{M} \dot{\vec{r}} \cdot \dot{\vec{r}}_{CM} \Big) \\ &+ \frac{1}{2} m_2 \Big(\big(\frac{m_1}{M}\big)^2 \Big|\dot{\vec{r}}\Big|^2 + \Big|\dot{\vec{r}}_{CM}\Big|^2 + \frac{2m_1}{M} \dot{\vec{r}} \cdot \dot{\vec{r}}_{CM} \Big) \\ &= \frac{1}{2}(m_1 + m_2)\Big|\dot{\vec{r}}\Big|_{CM}^2 + \frac{m_1 m_2^2 + m_2 m_1^2}{2(m_1 + m_2}\Big|\dot{\vec{r}}\Big|^2 \\ &= \frac{1}{2} M \Big|\dot{\vec{r}}\Big|_{CM}^2 + \frac{1}{2}\Big(\frac{m_2 m_2}{m_1 + m_2}\Big)\Big|\dot{\vec{r}}\Big|^2 \end{aligned}$$

Finally we define the *reduced mass* as

$$\mu \equiv \frac{m_2 m_2}{m_1 + m_2} \tag{3.78}$$

After all this change of variables, we end up with a Lagrangian of the form

$$\mathcal{L} = \left(\frac{1}{2}M\left|\dot{\vec{r}}\right|^2_{CM}\right) + \left(\frac{1}{2}\mu\left|\dot{\vec{r}}\right|^2 - V(|\vec{r}|)\right) \tag{3.79}$$

$$\equiv \quad \mathcal{L}_M \quad + \quad \mathcal{L}_\mu \tag{3.80}$$

Notice that we have also used the notion of isotropy about the center of mass to make V a central potential. As a result of these algebraic manipulations, we come to result where we can essentially split our Lagrangian into a *free-particle Lagrangian* $\mathcal{L}_M$ and a *central potential Lagrangian* $\mathcal{L}_\mu$.

This result allows us to analyze each Lagrangian separately, which makes our analysis of the central potential Lagrangian only dependent on **one** coordinate $r = |\vec{r}|$, making this a 1-D problem. It is also good to note that due to the rotational symmetry of our problem here, angular momentum conservation immediately follows from this by Noether's theorem (3.3.1).

From here, we continue our analysis of $\mathcal{L}_\mu$.

$$\mathcal{L}_\mu = \frac{1}{2}\mu\dot{r}^2 - V(r) \tag{3.81}$$

$$\Rightarrow \quad \frac{\partial \mathcal{L}_\mu}{\partial r} = F_r = \mu r\dot{\phi}^2 - \frac{\partial V(r)}{\partial r}, \quad \frac{\partial \mathcal{L}_\mu}{\partial \dot{r}} = p_r = \mu\dot{r} \tag{3.82}$$

$$\Rightarrow \quad L_z^2 = \left|\vec{r} \times \vec{p}\right|^2 = r^2 \cdot \left(\mu^2 r^2 \dot{\phi}^2\right) = \mu^2 r^4 \dot{\phi}^2 \tag{3.83}$$

$$\Rightarrow \quad \mu\ddot{r} = \mu r\dot{\phi}^2 - \frac{\partial V(r)}{\partial r} = \frac{L_z^2}{\mu r^3} - \frac{\partial V(r)}{\partial r} \tag{3.84}$$

$$\Rightarrow \quad \boxed{\ddot{r} = \frac{L_z^2}{\mu^2 r^3} - \frac{1}{\mu}\frac{\partial V(r)}{\partial r}} \tag{3.85}$$

Chapter 4

Orbits and Scattering

In the early 16 hundreds, a German astronomer by the name of Johannes Kepler modelled the orbits of celestial bodies around the sun, as a result he proposed 3 laws of planetary orbital motion around the sun. The last example in the previous chapter touches on 2-body problems which are reduced to single degree of freedom problems. We will be using this approach of reduction to derive and prove Kepler's laws.

§4.1 Motion of Celestial Bodies

First consider the energy of a celestial body orbiting the sun. To a good approximation, we can take the center of mass of the system to be at the center of the sun.The reason for this is shown below. First, we denote the mass of the sun as m_s and the mass of the orbiting body as m_p where $m_p \ll m_s$.

$$\vec{r}_{CM} = \frac{m_p\vec{r}_p + m_s\vec{r}_s}{m_p + m_s} \approx \frac{m_p}{m_s}\vec{r}_p + \vec{r}_s \approx \vec{r}_s \tag{4.1}$$

$$\mu = \frac{m_p m_s}{m_p + m_s} \approx \frac{m_p m_s}{m_s} = m_p \tag{4.2}$$

Now consider the total energy of a planet is orbital motion around the sun. It will have 3 components, a radial kinetic energy term, an angular momentum term and the central potential energy.

$$E = \frac{1}{2}\mu\dot{r}^2 + \frac{L_z^2}{2\mu r^2} + V(r) \tag{4.3}$$

Since the 2 last terms are solely dependent on r, L_z begin constant due to the isotropic potential leading to conservation of angular momentum, we can lump them into an *effective potential* $V_{\text{eff}}(r)$.

$$V_{\text{eff}}(r) \equiv \frac{L_z^2}{2\mu r^2} + V(r) \tag{4.4}$$

As such, the energy and Lagrangian of the system are now written as

$$E = \frac{1}{2}\mu\dot{r}^2 + V_{\text{eff}} \tag{4.5}$$

$$\mathcal{L} = \frac{1}{2}\mu\dot{r}^2 - V_{\text{eff}} \tag{4.6}$$

§4.2 Kepler's Laws of Orbital Motion

As earlier mentioned, Kepler was concerned with the motion of planets around the sun. As such, the potential we will be observing to cover Kepler's laws is one constructed from Newton's laws of gravitation. Even though this was **not** how Kepler himself arrived at his results (having done his work before Isaac Newton), it will be the easiest means for us to go about understanding his laws.

$$V(r) = -\frac{Gm_s m_p}{r} \approx -\frac{GM\mu}{r} \tag{4.7}$$

$$\Rightarrow \; V_{\text{eff}} = \frac{L_z^2}{2\mu r^2} - \frac{GM\mu}{r} \tag{4.8}$$

To get a better idea of things, we look at the graph of the effective potential against the radius of orbit.

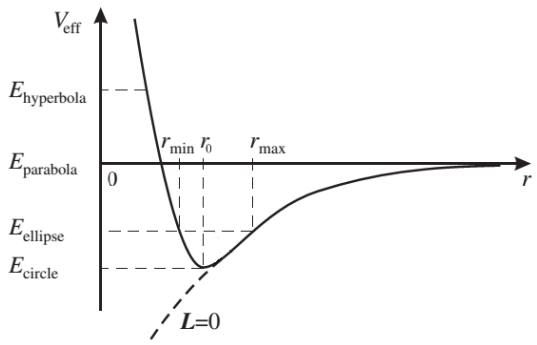

Figure 4.1: Effective Potential

The shape of the graph in figure 4.1 arises due to a combination of the $\frac{1}{r^2}$ and $-\frac{1}{r}$ terms in the effective potential. The result is the creation of a 'potential well' around $r = r_0$ with a minimum energy E_{circle}. This is what allows the planets to remain in a stable orbit around the sun! This immediately implies that for all planets in our solar system, the value of their total energy E has to be **less than** 0. With this, we can present and prove *Kepler's first law.*

Kepler's 1st Law

Orbits around the sun are elliptical with the sun at one of the 2 foci.

Proof. Consider stable orbits around the sun, meaning $E_{\text{circle}} \leq E < 0$. We also define the following parameters for convenience.

$$\lambda = \frac{L_z^2}{\mu}, \quad \alpha = GM\mu \tag{4.9}$$

$$\Rightarrow \; V_{\text{eff}}(r) = \frac{\lambda}{2r^2} - \frac{\alpha}{r} \tag{4.10}$$

$$\Rightarrow \; E = \frac{1}{2}m\dot{r}^2 + \frac{\lambda}{2r^2} - \frac{\alpha}{r} \tag{4.11}$$

$$\Rightarrow \; \dot{r} = \frac{dr}{dt} = \sqrt{\frac{2}{\mu}\Big(E + \frac{\alpha}{r} - \frac{\lambda}{2r^2}\Big)} \tag{4.12}$$

Also looking at the angular momentum of the system,

$$L_z = |\vec{r} \times \vec{p}| = \mu r^2 \dot{\phi} \tag{4.13}$$

$$\Rightarrow \; \frac{d\phi}{dt} = \frac{L_z}{\mu r^2}$$

$$\Rightarrow \; \int_0^{\phi_0} d\phi = \int_0^t \frac{L_z}{\mu r^2} dt$$

$$\Rightarrow \; \phi = \phi_0 + \int_{r(t=0)}^{r(t=t)} \frac{L_z}{\mu r^2} \frac{dr}{\sqrt{\frac{2}{\mu}\Big(E + \frac{\alpha}{r} - \frac{\lambda}{2r^2}\Big)}}$$

$$\Rightarrow \; \phi = \phi_0 + \int_{r(t=0)}^{r(t=t)} \frac{L_z dr}{\sqrt{2\mu\Big(E + \frac{\alpha}{r}\Big) - \frac{L_z^2}{r^2}}} \tag{4.14}$$

In order to solve the complicated integral in (4.14), we require to do a few substitutions along way.

$$u \equiv \frac{1}{r} \; \Rightarrow \quad dr = -\frac{du}{u^2}$$

$$\Rightarrow \; \phi - \phi_0 = \int_{u(t=0)}^{u(t=t)} \frac{L_z du}{u\sqrt{\frac{2\mu E}{u^2} + \frac{2\mu\alpha}{u} - L_z^2}}$$

$$= -\frac{L_z}{\sqrt{2\mu E}} \int_{u(t=0)}^{u(t=t)} \frac{du}{\sqrt{1 + \frac{\alpha}{E}u - \frac{L_z^2}{2\mu E}u^2}}$$

Let

$$a \equiv \frac{\alpha}{E}, \quad b \equiv \frac{L_z^2}{2\mu E}$$

$$\Rightarrow \; \phi - \phi_0 = -\frac{L_z}{\sqrt{2\mu E}} \int_{u(t=0)}^{u(t=t)} \frac{du}{\sqrt{1 + au - bu^2}}$$

$$= -\frac{L_z}{\sqrt{2\mu E}} \int_{u(t=0)}^{u(t=t)} \frac{du}{\sqrt{-\big(\sqrt{b}u - \frac{a}{2\sqrt{b}}\big)^2 + \frac{a^2}{4b} + 1}}$$

Let

$$x \equiv \frac{2bu-a}{2\sqrt{\frac{a^2}{4}+b}} \quad \Rightarrow \quad du = dx\sqrt{\frac{a^2}{4b^2}+\frac{1}{b}}$$

$$\Rightarrow \ \phi-\phi_0 = -\frac{L_z}{\sqrt{2\mu bE}}\int_{u(t=0)}^{u(t=t)}\frac{dx}{\sqrt{1-x^2}}$$

$$= -\frac{L_z}{\sqrt{2\mu bE}}\cos^{-1}(x) = -\frac{L_z}{\sqrt{2\mu E}}\sqrt{\frac{2\mu E}{L_z^2}}\cos^{-1}(x)$$

$$\Rightarrow \ \phi-\phi_0 = \cos^{-1}\left(\frac{2bu-a}{2\sqrt{\frac{a^2}{4}+b}}\right) = \cos^{-1}\left(\frac{\left(\frac{L_z^2}{\mu E}\right)\frac{1}{r}-\frac{\alpha}{E}}{2\sqrt{\frac{\alpha^2}{4E^2}+\frac{L_z^2}{2\mu E}}}\right) \tag{4.15}$$

Having this relation between ϕ and r, we now look back at figure 4.1. We see that if the energy of the system is exactly at the minimum of the effective potential, this implies there would be no radial variation from r_0 in the orbit making the orbit **circular**. To find the minimum energy,

$$\left.\frac{\partial V_{\text{eff}}}{\partial r}\right|_{r=r_0} = \left.-\frac{\lambda}{r^3}+\frac{\alpha}{r^2}\right|_{r=r_0} = 0 \tag{4.16}$$

$$\Rightarrow \ r_0 = \frac{\lambda}{\alpha} \tag{4.17}$$

$$\Rightarrow \ E_{\text{circle}} = \min\{V_{\text{eff}}\} = V_{\text{eff}}(r_0) = \frac{\alpha^2}{\lambda}\left(\frac{1}{2}-1\right) \tag{4.18}$$

We split the values within the parenthesis in equation (4.18) to indicate the kinetic and potential energy contributions to the total energy (positive repulsive term is due to kinetic energy, negative attractive term is due to potential energy). For convenience and consistency in notation, we also relabel E_{circle} as $E_0 < 0$. As such, we can further clean-up equation (4.15) to give

$$\cos(\phi-\phi_0) = \frac{\left(\frac{\mu E}{L_z}\right)}{\left(\frac{\mu E}{L_z}\right)} \times \frac{\left(\frac{L_z^2}{\mu E}\right)\frac{1}{r}-\frac{\alpha}{E}}{2\sqrt{\frac{\alpha^2}{4E^2}+\frac{L_z^2}{2\mu E}}}$$

$$= \frac{\frac{L_z}{r}-\frac{\mu\alpha}{L_z}}{\sqrt{2\mu E+\frac{\mu^2\alpha^2}{L_z^2}}}$$

$$= \sqrt{\frac{L_z^2}{2\mu\alpha^2}}\cdot\frac{\frac{\alpha}{r}-\frac{\mu\alpha^2}{L_z^2}}{\sqrt{E+\frac{\mu\alpha^2}{2L_z^2}}}$$

Since

$$E_0 = -\frac{\alpha^2}{2\lambda} = -\frac{\mu\alpha^2}{2L_z^2}$$
$$\Rightarrow \quad \cos(\phi - \phi_0) = \frac{1}{2\sqrt{-E_0}} \cdot \frac{\frac{\alpha}{r} + 2E_0}{\sqrt{E - E_0}}$$
$$\Rightarrow \quad \cos(\phi - \phi_0) = \frac{1 + \frac{\alpha}{2E_0 r}}{\sqrt{1 - \frac{E}{E_0}}}$$

We then define the denominator as a new quantity,

$$e \equiv \sqrt{1 - \frac{E}{E_0}} \tag{4.19}$$
$$\Rightarrow \quad \boxed{e \cdot \cos(\phi - \phi_0) = 1 - \frac{r_0}{r}} \tag{4.20}$$

Which is indeed the equation of an ellipse! □

We also see from equation (4.15) that the maximum and minimum radii can be attained by taking the peak values of a cosine function.

$$\Rightarrow r_{\min} = \frac{r_0}{1+e}, \quad r_{\max} = \frac{r_0}{1-e} \tag{4.21}$$

Additionally, the last quantity that was defined in equation (4.19) before the final result is known as the *eccentricity* of the orbit.

Definition 4.2.1. Eccentricity: *The eccentricity e of an orbit is its degree of deviation from a circular path, where $e = 0$ denotes a perfectly circular orbit and larger values indicate larger deviations. All bound orbits have $e < 1$.*

We now continue down the list in Kepler's series of laws.

Keplers 2nd Law

A line segment connecting the Sun to its orbiting planet sweeps out equal area in equal time intervals.

Proof. From the statement of Kepler's law,

$$\frac{dA}{dt} = \text{constant}$$
$$\Rightarrow \quad \frac{dA}{dt} = \frac{1}{2} r(\vec{v}\phi) = \frac{1}{2} r^2 \dot{\phi} = \frac{L_z}{2\mu} = \text{constant}$$
$$\Rightarrow \quad L_z = \text{constant} \tag{4.22}$$

Hence we see that this law is equivalent to the conservation law of angular momentum due to central, isotropic potentials which we have already proven in chapter 3. □

Finally, we present and prove Kepler's third and final law.

Kepler's 3rd Law

The square of the orbital period is proportional to the cube of the semi-major axis of orbit.

Proof. From Kepler's second law, we have that

$$\frac{L_z}{2\mu} = \frac{dA}{dt} \tag{4.23}$$

with all of the terms on the right being constant. Hence performing an integral with respect to time on both sides, we get

$$\int_0^\tau \frac{L_z}{2\mu} dt = \int_0^\tau \frac{dA}{dt} dt \tag{4.24}$$

$$\Rightarrow \; A = \frac{L_z}{2\mu}\tau = \pi a b \tag{4.25}$$

Where a is the length of the semi-major axis and b is the length of the semi-minor axis. By the geometry of an ellipse, we get that

$$a = \frac{r_0}{1-e^2} = -\frac{\alpha}{2E} \tag{4.26}$$

$$b = \frac{r_0}{\sqrt{1-e^2}} = \frac{L_z}{\sqrt{-2\mu E}} = \sqrt{r_0 a} \tag{4.27}$$

We further perform manipulations on b to give

$$\begin{aligned} b \quad &= \frac{L_z}{\sqrt{-2\mu E}} = \frac{L_z}{2\mu}\sqrt{2\mu - E} \\ &= \frac{L_z}{2\mu}\sqrt{\frac{4a\mu}{\alpha}} = \frac{L_z}{\mu}\sqrt{\frac{a\mu}{\alpha}} \\ \Rightarrow \; \frac{L_z}{2\mu}\tau &= \pi a \frac{L_z}{\mu}\sqrt{\frac{a\mu}{\alpha}} \\ \Rightarrow \; \tau \quad &= 2\pi a^{\frac{3}{2}}\sqrt{\frac{\mu}{\alpha}} \end{aligned}$$

Hence we see that after squaring both sides of the final equation above, we indeed arrive at the result that

$$\boxed{\tau^2 \propto a^3} \tag{4.28}$$

□

We now move out of the realm of bound orbits and look at energies $E \geq 0$. What essentially is happening here is that the planet has too large a kinetic to be held in by the gravity of the

Sun. We first look at the case where ($\underline{E = 0}$).

$$
\begin{aligned}
& e(E) = \sqrt{1 - \frac{E}{E_0}} \\
\Rightarrow \quad & e(E)|_{E=0} = \sqrt{1 - \frac{E}{E_0}}\bigg|_{E=0} = 1 && (4.29) \\
\Rightarrow \quad & \cos(\phi - \phi_0) = 1 - \frac{r_0}{r} && (4.30) \\
\Rightarrow \quad & r_{\min} = \frac{r_0}{2}, \quad r_{\max} \to \infty && (4.31)
\end{aligned}
$$

This occurs when $T = -V_{\text{eff}}$ and as we see from equation (4.30), this results in a *parabolic path* around the Sun. This is confirmed because we see that the maximum orbit radius $r_{\max}$ goes off to infinity. We now look at the case where ($\underline{E > 0}$).

$$
\begin{aligned}
& e(E) = \sqrt{1 - \frac{E}{E_0}} \\
\Rightarrow \quad & e(E)|_{E>0} = \sqrt{1 - \frac{E}{E_0}}\bigg|_{E>0} < 1 && (4.32) \\
\Rightarrow \quad & \cos(\phi - \phi_0) < 1 - \frac{r_0}{r} && (4.33)
\end{aligned}
$$

This occurs when $T > -V_{\text{eff}}$ and as we see from equation (4.33), this results in a *hyperbolic path* around the Sun. Note here that for hyperbolic orbits, the range of ϕ is restricted to

$$
\boxed{\phi_\infty = \cos^{-1}(-\frac{1}{e})} \tag{4.34}
$$

where ϕ_∞ is the acute angle between the asymptote and the x-axis. We will now work through an effective potential example problem below.

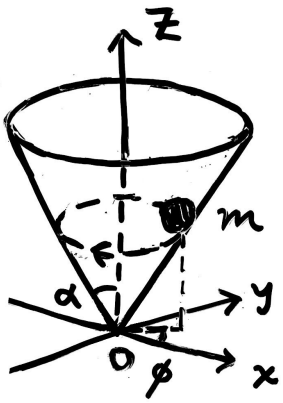

Figure 4.2: Mass in Cone System

Example

Consider a inverted cone with its tip at the origin of our Cartesian coordinate system. There is a ball of mass m that moves along the inner surface of the cone and experiences downward acceleration due to gravity long the z-axis. The cone has an angle of inclination α from the z-axis. Refer to figure 4.2 for a visualization of the problem.

Our coordinates are thus given as,

$$x = r' \sin(\alpha)\cos(\phi) = r\cos(\phi) \tag{4.35}$$
$$y = r' \sin(\alpha)\sin(\phi) = r\sin(\phi) \tag{4.36}$$
$$z = r' \cos(\alpha) = r\cot(\phi) \tag{4.37}$$
$$r \equiv r' \sin(\alpha) \tag{4.38}$$

We can now construct the kinetic and potential energies of our system as follows.

$$T = \frac{1}{2}m(\dot{x}^2 + \dot{y}^2 + \dot{z}^2) = \frac{1}{2}m\big(\dot{r}^2 + r^2\dot{\phi}^2 + \dot{r}^2\cot^2(\alpha)\big) \tag{4.39}$$
$$V = mgz = mgr\cot(\alpha) \tag{4.40}$$
$$\Rightarrow \quad \mathcal{L} = T - V = \frac{1}{2}m\big(r^2\dot{\phi}^2 + \dot{r}^2\csc^2(\alpha)\big) - mgr\cot(\alpha) \tag{4.41}$$
$$\Rightarrow \quad \frac{\partial \mathcal{L}}{\partial \phi} = 0 \tag{4.42}$$

Notice here that our potential is positive but we still get bound orbits. This is due the 0 potential energy reference we have defined for this problem (not at infinity but at the bottom of the cone).

From (4.42), we see that angular momentum of the system is conserved.

$$\Rightarrow \quad L_z = mr^2\dot{\phi} = \text{constant} \tag{4.43}$$
$$\Rightarrow \quad E = \frac{1}{2}m\dot{r}^2\csc^2(\alpha) + \Big(\frac{L_z}{2mr^2} + mgr\cot(\alpha)\Big)$$
$$m' \equiv m\csc^2(\alpha), \quad \boxed{V_{\text{eff}}(r) \equiv \frac{L_z}{2mr^2} + mgr\cot(\alpha)} \tag{4.44}$$
$$\Rightarrow \quad E = \frac{1}{2}m'\dot{r}^2 + V_{\text{eff}}(r) \tag{4.45}$$

From here, we can look for the *turning points* of the system. These occur at points where $\dot{r} = 0$. As a result, the kinetic energy term vanishes and we get

$$E = V_{\text{eff}}(r) = \frac{L_z}{2mr^2} + mgr\cot(\alpha) \tag{4.46}$$
$$\Rightarrow \quad 2m^2 g\cot(\alpha)r^3 - 2Emr^2 + L_z^2 = 0 \tag{4.47}$$

Even though this is a cubic equation, there are only 2 physically feasible turning point solutions as would be expected. Note that if the energy of the mass falls below the effective potential, it would cease to orbit and simply fall to the bottom of the cone (turning points would be imaginary).

§4.3 Laplace-Runge-Lenz Vector

Continuing our analysis of central potentials in the context of celestial bodies, we have thus far seen 4 quantities being conserved as a result of an isotropic potential. These are namely, energy and the 3 components of angular momentum. However, we recall from chapter 3 that for a central potential in 3-D (3 degrees of freedom), there are actually $2(3) - 1 = 5$ integrals of motion! It turns out that the last conserved quantity resides in what is known as the *Laplace-Runge-Lenz vector*.

Consider a central potential $V(r)$, the force associated to the potential is then given by $\vec{F} = -\boldsymbol{\nabla} V(r) = f(r)\hat{r} = \dot{\vec{p}}(r)$. Now consider the following quantity.

$$\begin{aligned}\dot{\vec{p}} \times \vec{L} &= \frac{f(r)}{r}\vec{r} \times (\vec{r} \times m\dot{\vec{r}}) \\ &= \frac{mf(r)}{r}\big(\vec{r} \times (\vec{r} \times \dot{\vec{r}})\big) \\ &= \frac{mf(r)}{r}\big(\vec{r}(\vec{r} \cdot \dot{\vec{r}}) - r^2\dot{\vec{r}}\big) \end{aligned} \tag{4.48}$$

Then from the total time derivative identity,

$$\begin{aligned}\frac{d}{dt}(\vec{r} \cdot \vec{r}) &= \frac{d}{dt}(r^2) \\ &= 2\dot{\vec{r}} \cdot \vec{r} = 2r\dot{r}\end{aligned} \tag{4.49}$$

which reduces (4.48) to

$$\dot{\vec{p}} \times \vec{L} = \frac{mf(r)}{r}\big(\vec{r}(\dot{r}r) - r^2\dot{\vec{r}}\big) \tag{4.50}$$

$$\frac{d}{dt}(\vec{p} \times \vec{L}) = \dot{\vec{p}} \times \vec{L} + \vec{p} \times \dot{\vec{L}} = \dot{\vec{p}} \times \vec{L} \tag{4.51}$$

$$\Rightarrow \quad \frac{d}{dt}(\vec{p} \times \vec{L}) = -mf(r)r^2\Big(\frac{\dot{\vec{r}}}{r} - \frac{\vec{r}\dot{r}}{r^2}\Big) = -mf(r)r^2\Big(\frac{d}{dt}\hat{r}\Big) \tag{4.52}$$

Substituting the Newtonian gravity force law,

$$f(r) = -\frac{k}{r^2} \tag{4.53}$$

$$\Rightarrow \quad \frac{d}{dt}(\vec{p} \times \vec{L}) = mk\frac{d}{dt}(\hat{r}) \tag{4.54}$$

Looking at this result, allow us to then define a quantity

$$\boxed{\vec{A} \equiv \vec{p} \times \vec{L} - \frac{mk}{r}\vec{r}} \tag{4.55}$$

$$\Rightarrow \quad \frac{d\vec{A}}{dt} = mk\frac{d}{dt}(\hat{r}) - mk\frac{d}{dt}\Big(\frac{\vec{r}}{r}\Big) = 0 \tag{4.56}$$

Showing that this newly defined quantity in (4.55) is another conserved quantity in central potential systems! This quantity is indeed, as mentioned, the Laplace-Runge-Lenz vector. A visulization of this vector can be seen in figure 4.3. In this plot 4.3, we are setting up the orbit

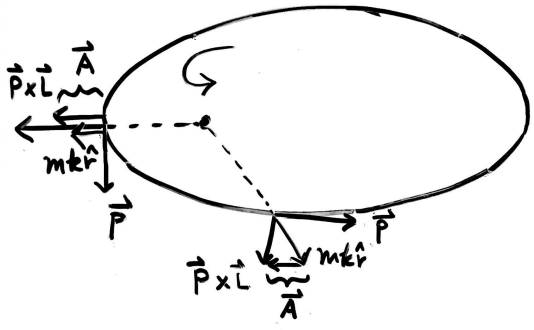

Figure 4.3: Laplace-Runge-Lenz Vector Visualization

such that it only moves in the x,y-plane, where $\vec{L}$ points toward the reader. We can see just from the visual geometry of the figure that no matter where the planet is in its orbit, the direction is $\vec{A}$ is preserved. Its magnitude is preserved due to the scaling factors attached to its constituent vectors in the definition.

Things so far all seem well and good but an astute observer may notice that this new conserved quantity is a vector quantity, which means there are implicitly 3 new conserved quantities! This means that our system must then have 2 constraints so that we end up with exactly the 5 integrals of motion that we would expect. Just looking at the figure, it is likely that one can already see what one of these constrains would be.

1st *Constraint*:

$\vec{A}$ **must** lie in the plane of the orbit. This can easily be shown by taking the dot product of $\vec{A}$ with $\vec{L}$, since $\vec{L}$ is always orthogonal to the orbit plane.

$$\begin{aligned}\vec{A}\cdot\vec{L} &= \left(\vec{p}\times\vec{L} - \frac{mk\vec{r}}{r}\right)\cdot(\vec{r}\times\vec{p}) \\ &= \left(\vec{p}\times(\vec{r}\times\vec{p})\right)\cdot(\vec{r}\times\vec{p}) - \frac{mk}{r}\vec{r}\cdot(\vec{r}\times\vec{p}) = 0\end{aligned} \tag{4.57}$$

which implies that $\vec{A}$ is always orthogonal to $\vec{L}$ and hence, indeed lies in the plane of orbit.

2nd *Constraint*:

This next constraint is a little more involved and basically says that a particular function $f(\vec{A})$ is exactly equivalent to the orbit equation (4.20). To show this, we first define the angle between the vectors $\vec{A}$ and $\vec{r}$ as θ. It then follows that

$$\begin{aligned}\vec{A}\cdot\vec{r} &= \left|\vec{A}\right||\vec{r}|\cos(\theta) = \left(\vec{p}\times\vec{L} - mk\frac{\vec{r}}{r}\right)\cdot\vec{r} \\ &= \vec{L}\cdot(\vec{r}\times\vec{p}) - mk\frac{\vec{r}\cdot\vec{r}}{r} = \left|\vec{L}\right|^2 - mkr\end{aligned} \tag{4.58}$$

Rearranging this result, we get

$$\frac{1}{r} = \frac{mk}{L^2}\left(1 + \frac{A}{mk}\cos(\theta)\right) \tag{4.59}$$

which then if we set $e \equiv \frac{A}{mk}$ and $r_0 \equiv \frac{L^2}{mk}$, we get exactly the elliptical orbit equation

$$\frac{1}{r} = \frac{1}{r_0}\big(1 - e\cos(\theta)\big) \tag{4.60}$$

We can then further extend this by recalling our definition of the eccentricity,

$$e = \sqrt{1 + \frac{2EL^2}{mk^2}} \tag{4.61}$$

$$\Rightarrow \quad A^2 = m^2k^2 + 2mEL^2 \tag{4.62}$$

eliminating one more constraint we leaves us with 5 just as we expected.

§4.4 Scattering

We now look at problems with *repulsive potentials*. A repulsive potential would result in **no** possibility of any orbits but would cause the incoming object to scatter off it. One such example of a repulsive potential would be $V(r) = \frac{\alpha}{r}$. In general, the set-up of a scattering problem is as follows.

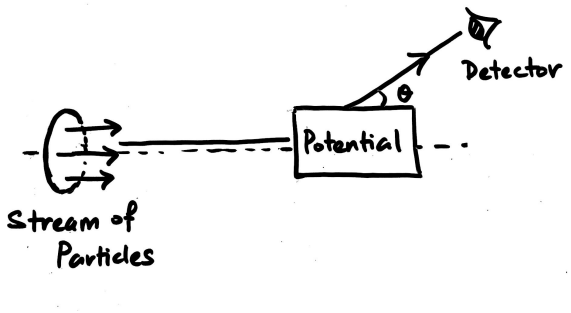

Figure 4.4: Scattering Set-Up

Generally when we speak of scattering problems, we think of particles scattering off a particular potential. As such figure 4.4 illustrates this. The angle θ at which the particle is deflected from its original trajectory is known as the *scattering angle*. An example of a specific scattering system is protons scattering off gold atoms.

In scattering experiments, the aim is to attain information about a given potential by looking at how particles scatter off it. As such, we look for a quantity known as the *scattering cross-section*.

Definition 4.4.1. Scattering Cross-Section: *The scattering cross-section σ is defined as the ratio of the flow of scattered particles and the incident particle flux.*

$$\sigma = \frac{(\text{Number of scattered particles})/(\text{time})}{(\text{Number of incident particles})/(\text{Area} \times \text{time})} \tag{4.63}$$

Simply put, this gives the cross-sectional area of the region where incident particles would scatter off the potential. Before we continue to derive a mathematical expression that can be used to solve for σ, we look at what is known as the *solid angle*.

Definition 4.4.2. Solid Angle: *The solid angle is the 2 dimensional angle defined in 3 dimensional space that an object subtends at a point. The units of a solid angle is called steradians.*

To get a visual understanding of this, consider the differential solid angle subtended to 2 different surfaces with areas A_1 and A_2 at a distant R_1 and R_2 away from it respectively. This is illustrated in figure 4.5.

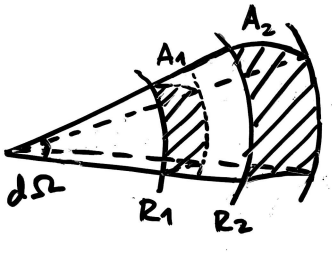

Figure 4.5: Differential Solid Angle

We then see from geometrical arguments that

$$d\Omega = \frac{A_1}{R_1^2} = \frac{A_2}{R_2^2} \tag{4.64}$$

Considering now spherical coordinates, and as per usual taking ϕ to be the azimuthal angle and θ to be the angle of inclination,

$$\Rightarrow \boxed{d\Omega = \sin(\theta) d\theta d\phi}, \quad \phi \in [0, 2\pi], \theta \in [0, \pi] \tag{4.65}$$

Now consider the quantity

$$\frac{d\sigma}{d\Omega} = \frac{(\text{Number of scattered particles})/(\text{time} \times \text{steradians})}{(\text{Incident flux})} \tag{4.66}$$

From this definition, we get

$$\sigma = \int_0^{2\pi} d\phi \int_0^{2\pi} d\theta \sin(\theta) \frac{d\sigma}{d\Omega} \tag{4.67}$$

In the case of our scattering set-up, we define our θ and ϕ angles as illustrated below.

As seen in figure 4.6, we have also defined a differential ring of thickness $d\theta$ on the scattered plane and a differential ring off thickness db on the incident stream plane where b is the distance from the central axis and known as the *impact parameter*.

Also, if our target potential is a central potential $\big(V(\vec{r}_1, \vec{r}_2) = V(|\vec{r}_1 - \vec{r}_2|) = V(r)\big)$, we get that $d\sigma/d\Omega$ should **only** depend on θ and **not** ϕ. As such, the ϕ dependence can be integrated

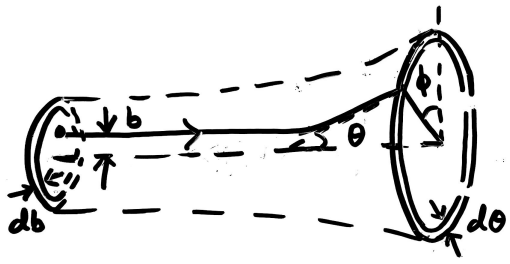

Figure 4.6: Scattering Coordinates

independent of the function of θ to pull out a constant factor of 2π (refer to equation (4.67)). Also this makes our differential solid angle for all intends and purposes effectively

$$d\Omega = \sin(\theta)d\theta \int_0^{2\pi} d\phi = 2\pi \sin(\theta)d\theta \tag{4.68}$$

Now, using these new quantities, we can construct the *fraction of scattered particles* into a differential ring of thickness $d\theta$ in the scattered plane as

$$\begin{aligned} f_{\text{scatt}} &= \frac{N_{\text{scatt}}}{N_{\text{inc}}} = \frac{1}{a}\Big(\frac{d\sigma}{d\Omega}\Big)d\Omega \qquad (4.69)\\ &= \frac{1}{a}\Big(\frac{d\sigma}{d\Omega}\Big)\Big(\frac{d\Omega}{d\theta}\Big)d\theta \\ &= \frac{1}{a}\Big(\frac{d\sigma}{d\Omega}\Big)\Big(2\pi \sin(\theta)\Big)d\theta \qquad (4.70)\end{aligned}$$

where a is the total cross-sectional area of the incident stream of particles.

Similarly, we can construct the fraction of scattered particles that come from a differential ring of thickness db on the incident stream plane.

$$f_{\text{scatt}} = \frac{N_{\text{scatt}}}{N_{\text{inc}}} = \frac{1}{a}(2\pi b)db \tag{4.71}$$

Then by *particle number conservation*, we can put the 2 equations together to give

$$\begin{aligned} &\frac{1}{a}\Big(\frac{d\sigma}{d\Omega}\Big)\Big(2\pi \sin(\theta)\Big)d\theta = \frac{1}{a}(2\pi b)db \\ \Rightarrow\ &\boxed{\frac{d\sigma}{d\Omega} = \frac{b}{\sin(\theta)}\left|\frac{db}{d\theta}\right|} \qquad (4.72)\end{aligned}$$

where we take the absolute value of $\frac{db}{d\theta}$ because we are not concerned with the parity of the quantity here. To show a utilization of this result, we will work through an example problem.

Example

Consider a scattering experiment where a stream of incident particles are scattered off a hard sphere of radius R. Here, we denote the scattering angle as θ and the distance from

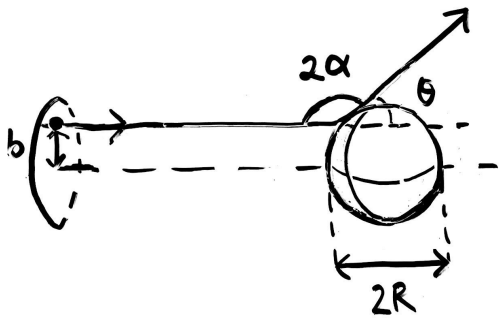

Figure 4.7: Scattering Off Hard Sphere

the central axis to be b. This is illustrated in figure 4.7. Before we begin to solve this analytically, let us walk through an intuitive argument and see if things work out as our intuition would suggest.

Intuitive Argument:
Seeing that the goal here is to determine the scattering cross-section σ, it is fair to guess that the scattering cross-section we would get from this is πR^2. The idea behind this is that an incident particle would be scattered **only** if it collides with the hard sphere. Hence, this only occurs if the incident particle is incident within the cross-section of the sphere which is exactly πR^2.

Rigorous Derivation:
Allow us to now see if our intuition indeed corresponds to the rigorous mathematical result. First we look at a couple of geometrical relations.

$$\theta = \pi - 2\alpha \tag{4.73}$$

$$\Rightarrow \; b = R\sin(\alpha) = R\sin\left(\frac{\pi-\theta}{2}\right)$$

$$\Rightarrow \; b = R\cos\left(\frac{\theta}{2}\right) \tag{4.74}$$

Then using equation (4.72) derived earlier, we substitute the geometrical quantities above to get

$$\frac{d\sigma}{d\Omega} = \frac{R\cos\left(\frac{\theta}{2}\right)}{\sin(\theta)}\left|-\frac{R}{2}\sin\left(\frac{\theta}{2}\right)\right| = \frac{R^2}{4} \tag{4.75}$$

Integrating with respect to Ω, we get

$$\begin{aligned}\sigma &= 2\pi\int_0^{\pi} d\theta\,\sin(\theta)\frac{R^2}{4} \\ &= \frac{R^2}{4}(4\pi) = \pi R^2\end{aligned} \tag{4.76}$$

Indeed arriving at the result we proposed! Note that the 2π in front of the integral above came from the ϕ being integrated.

§4.4.1 Rutherford Scattering

We now explore a famous scattering experiment explained by Ernest Rutherford in 1911. Rutherford scattering is the elastic scattering of charged particles by off a potential due to Coulombic interactions. The set-up is identical to that illustrated in figure 4.4. The potential here is a repulsive central potential and can be modelled as

$$V(r) = \frac{\alpha}{r} \tag{4.77}$$

where the α defined here is not exactly $GM\mu$ as previously defined, but can be taken to be an analogous quantity (an electrostatic version).

We normally take particles as coming from far away (from ∞) and hence are incoming with a velocity v_∞. Because we are dealing here with an isotropic potential, angular momentum of the incident particles will be conserved and remain constant throughout the scattering experiment. The angular momentum and incident energy of the incoming particles are thus given as

$$L = mbv_\infty \tag{4.78}$$

$$E_{\text{incident}} = \frac{1}{2}mv_\infty^2 \tag{4.79}$$

where again, b is the distance from the central scattering axis. Taking the results from Kepler's laws, we have

$$\cos(\phi_0) = \frac{1}{e}, \quad e = \sqrt{1 + \frac{2EL^2}{m\alpha^2}} \tag{4.80}$$

where e is the eccentricity, $r_0 = 0$ since there is no possible circular orbit and the angle ϕ_0 is defined as equivalent to α in figure 4.7. We use ϕ_0 instead to avoid confusion with the α defined in the potential.

Then by the geometry of the problem,

$$\sin\left(\frac{\theta}{2}\right) = \sin\left(\frac{\pi}{2} - \alpha\right) = \cos(\alpha) = \frac{1}{e} \tag{4.81}$$

$$= \frac{1}{\sqrt{1 + \frac{2EL^2}{m\alpha^2}}} = \frac{1}{\sqrt{1 + 2E\frac{m^2b^2v_\infty^2}{m\alpha^2}}} \tag{4.82}$$

$$= \frac{1}{\sqrt{1 + \left(\frac{2Eb}{\alpha}\right)^2}} \tag{4.83}$$

$$\Rightarrow \left(\frac{2Eb}{\alpha}\right)^2 = \frac{1}{\sin^2(\frac{\theta}{2})} - 1 = \frac{1}{\tan^2(\frac{\theta}{2})} \tag{4.84}$$

which simplifies to

$$\boxed{b = \frac{\alpha}{2E\tan\left(\frac{\theta}{2}\right)}} \tag{4.85}$$

$$\Rightarrow \left|\frac{db}{d\theta}\right| = \frac{\alpha}{4E} \cdot \frac{1}{\sin^2\left(\frac{\theta}{2}\right)} \tag{4.86}$$

$$\Rightarrow \frac{d\sigma}{d\Omega} = \frac{\alpha}{2E\sin(\theta)\tan\left(\frac{\theta}{2}\right)}\left(\frac{\alpha}{4E} \cdot \frac{1}{\sin^2\left(\frac{\theta}{2}\right)}\right) \tag{4.87}$$

$$\Rightarrow \frac{d\sigma}{d\Omega} = \left(\frac{\alpha}{4E\sin^2\left(\frac{\theta}{2}\right)}\right)^2 \tag{4.88}$$

Looking at equation (4.88), we see that if we look at a plot of $\frac{d\sigma}{d\Omega}$ against θ, there is a vertical asymptote at $\theta = 0$. What this physically means is that there is a definite chance of scattering off the potential when we have a particle incident along the scattering axis. Continuing our analysis, we integrate equation (4.88). Here we cannot integrate from $\theta = 0$ since the function value diverges there and so does the integral. So we integrate from an arbitrary angle $\theta_0 > 0$.

$$\sigma(\theta > \theta_0) = 2\pi \int_{\theta_0}^{\pi} d\theta \sin(\theta)\left(\frac{\alpha}{4E\sin^2\left(\frac{\theta}{2}\right)}\right)^2 \tag{4.89}$$

$$= \left(\frac{\alpha}{4E}\right)^2\left(\frac{4\pi}{\tan^2\left(\frac{\theta_0}{2}\right)}\right) = \pi\left(\frac{\alpha}{2E\tan\left(\frac{\theta_0}{2}\right)}\right)^2 \tag{4.90}$$

Then if we look at the result from (4.85), we arrive at

$$\boxed{\sigma(\theta > \theta_0) = \pi b^2(\theta_0)} \tag{4.91}$$

This result looks very similar to what we got for scattering off a hard sphere, but take note that the σ for Rutherford scattering depends on the distance from the central axis of the incident particles (b) and **not** a length parameter of the potential itself.

Chapter 5

Oscillations and Perturbation Theory

Oscillations occur in systems where the potential energy of the system contains local minima. Now consider a 1-D system with an arbitrary potential $V(q)$ having a local minima at q_0 as shown in figure 5.1.

Then taking the Taylor expansion of this potential around q_0, we get

$$V(q) = V(q_0) + \frac{\partial V(q_0)}{\partial q}(q - q_0) + \frac{1}{2}\frac{\partial^2 V(q_0)}{\partial q^2}(q - q_0)^2 + ... \tag{5.1}$$

From here, if we are only concerned with small movements around q_0, this means that $(q - q_0) << 1$, allowing us to ignore $\mathcal{O}((q-q_0)^3)$ terms and higher. As such, our potential is effectively just

$$V(q) = V(q_0) + \frac{\partial V(q_0)}{\partial q}(q - q_0) + \frac{1}{2}\frac{\partial^2 V(q_0)}{\partial q^2}(q - q_0)^2 \tag{5.2}$$

This is all well and good, but how do we find the point of *stable equilibrium* q_0 in the first place? It turns out that there are *2 conditions* that stable equilibrium points must satisfy and are the same conditions that follow from the definition of a local minima.

1. The **first order** stationary point condition, $\frac{\partial V(q_0)}{\partial q} = 0$
2. The **second order** positivity condition of local minima, $\frac{\partial^2 V(q_0)}{\partial q^2} \geq 0$

Notice how the first condition implies that the net force at q_0 is 0, which gives physical meaning to this condition. Enforcing these conditions on our potential around q_0, we get that

$$V(q) = V(q_0) + \frac{1}{2}\frac{\partial^2 V(q_0)}{\partial q^2}(q - q_0)^2 \tag{5.3}$$

Then if we simply shift our potential energy function such that this local minimum is the zero point, we can define the following quantites to get

$$x \equiv q - q_0, \quad k \equiv \frac{\partial^2 V(q_0)}{\partial q^2} \tag{5.4}$$

$$\Rightarrow \quad V(x) = \frac{1}{2}kx^2 \tag{5.5}$$

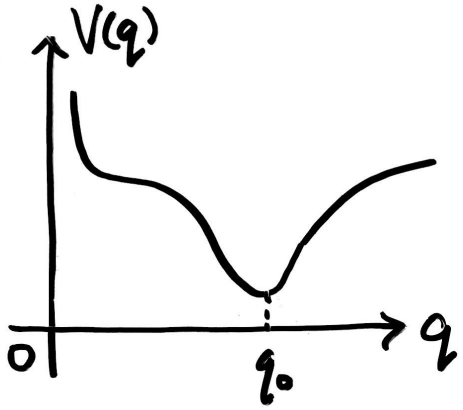

Figure 5.1: Arbitrary Potential with Minima at q_0

which is known as a *simple harmonic oscillator* potential. We will be studying these potentials in this chapter and how small perturbations away from the stable equilibrium affect the trajectories of objects. We now look at a use of this approximation in the examples below.

Example

Consider a satellite off mass m orbiting the Earth. We can take the mass of the satellite to be much less than that for the Earth, allowing us to treat this as a central potential problem (masses are taken to be μ and M). Utilizing Kepler's laws of orbital motion, we get that the effective potential of the system is

$$V_{\text{eff}}(r) = \frac{\lambda}{2r^2} - \frac{\alpha}{r} \tag{5.6}$$

where the plot of this potential is given in figure 4.1 with the minimum ocurring at $r_0 = \frac{\lambda}{\alpha}$. Zooming into the region around r_0, we can perform a Taylor expansion of the effective potential and enforce the stable equilibrium properties to get

$$\begin{aligned} V(r) &= V(r_0) + \left(\frac{\alpha}{r^2} - \frac{\lambda}{r^3}\right)\Big|_{r=r_0}(r-r_0) + \left(\frac{3\lambda}{r^4} - \frac{2\alpha}{r^2}\right)\Big|_{r=r_0}(r-r_0)^2 \\ &= V(r_0) + 0 + \frac{1}{2}\frac{\alpha^4}{\lambda^3}(r-r_0)^2 \end{aligned} \tag{5.7}$$

then defining

$$x \equiv r - r_0, \quad k \equiv \frac{\alpha^4}{\lambda^3} \tag{5.8}$$

$$\Rightarrow \; V(x) = \frac{1}{2}kx^2 \tag{5.9}$$

which shows that the potential around r_0 can be taken as that of a simple harmonic oscillator! Now consider that at some time, the circularly orbiting satellite briefly fires

it's thrusters radially toward the Earth. This cause satellite to a small *perturbation* away from its original circular trajectory. From here, we get the equation of motion and its solution from the simple harmonic oscillator potential as follows,

$$m\ddot{x}(t) + kx(t) = 0 \tag{5.10}$$

$$\Rightarrow \; x(t) = a\cos\left(t\sqrt{\frac{k}{m}} + \phi\right) \tag{5.11}$$

where a is the amplitude of oscillation (small) and ϕ is the phase offset from a perfect sine function.

$$\Rightarrow \; r(t) = r_0 + a\cos\left(t\sqrt{\frac{k}{m}} + \phi\right) \tag{5.12}$$

$$\omega \equiv \sqrt{\frac{k}{m}} = \sqrt{\frac{\alpha^4}{\mu\lambda^3}} \tag{5.13}$$

$$\Rightarrow \; T_{\text{oscillation}} = \frac{2\pi}{\omega} = 2\pi\sqrt{\frac{\mu\lambda^3}{\alpha^4}} \tag{5.14}$$

Comparing the result in (5.14) to the circular orbital period of the satellite which is the same algebraic result as the period of orbit of the Earth around the Sun,

$$\begin{aligned} \Rightarrow T_{\text{orbit}} &= \pi\alpha\sqrt{\frac{\mu}{-2E^3}} = \pi\alpha\sqrt{\frac{\mu}{-2(\alpha^2/2\lambda)^3}} \\ &= 2\pi\sqrt{\frac{\mu\lambda^3}{\alpha^4}} = T_{\text{oscillation}} \end{aligned} \tag{5.15}$$

Hence we see that the period of radial oscillations is exactly equal to the period of orbit, making the orbit be perturbed into a slightly elliptical shape.

Example

Consider a spring pendulum system which is constrained from below by a taught string of length R. The spring has a spring constant k and equilibrium length l_0 when stretched by mass m. Refer to figure 5.2 for a visualization. We will take the base of the string to be the origin and for simplicity, we will assume that the natural length of the spring is 0. We can then construct the Lagrangian as

$$\mathcal{L} = T - V \tag{5.16}$$

$$T = \frac{1}{2}m(R\dot{\phi})^2, \quad V = \frac{1}{2}k\big(R^2\sin^2(\phi) + (l_0 + R(1-\cos(\phi))^2\big) + mgR\cos(\phi) \tag{5.17}$$

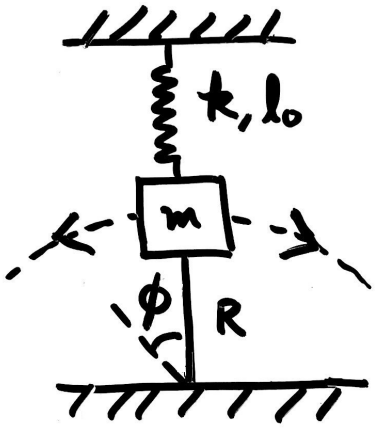

Figure 5.2: Constrained Spring Pendulum

Then using the small angle approximation on ϕ taking terms up to second order, we get

$$\begin{aligned} V &\approx \frac{1}{2}k\big(R^2\phi^2 + (l_0 + R(\frac{1}{2}\phi^2))^2\big) + mgR(1 - \frac{1}{2}\phi^2) \\ &\approx \frac{1}{2}k\big(R^2\phi^2 + l_0^2 + Rl_0\phi^2\big) + mgR(1 - \frac{1}{2}\phi^2) \\ &= \frac{1}{2}\big(kR^2 + kl_0R - mgR\big)\phi^2 + \big(\frac{1}{2}kl_0^2 + mgR\big) \end{aligned} \tag{5.18}$$

Notice that the second term in the result above is constant and thus can be dropped in the Lagrangian. The Lagrangian is thus

$$\mathcal{L} = \frac{1}{2}mR^2\dot{\phi}^2 - \frac{1}{2}\tilde{k}R^2\phi^2, \quad \tilde{k} \equiv k + \frac{kl_0}{R} - \frac{mg}{R} \tag{5.19}$$

where $\tilde{k}$ is the *effective spring constant*. Then since the angular frequency of oscillators are given as the square root of the spring constant over the oscillating mass,

$$\Rightarrow \boxed{\omega^2 = \frac{k}{m}\big(1 + \frac{l_0}{R}\big) - \frac{g}{R}} \tag{5.20}$$

Since this is supposed to be a squared quantity, the value on the right has to be positive. But in the event that $\frac{g}{R} > \frac{k}{m}\big(1 + \frac{l_0}{R}\big)$, we see that ω becomes a complex value. What this physically means is that the string no longer remains taught as gravity overcomes the lifting force of the spring.

§5.1 Driven Oscillations

We now look at perturbed systems with an external time-dependent force which drives the mass. We add this time-dependent force to the potential energy which becomes

$$V(x,t) = \frac{1}{2}kx^2 - xF(t) \tag{5.21}$$

$$\Rightarrow \; F_x = -\frac{\partial V(x,t)}{\partial x} = -kx + F(t) \tag{5.22}$$

$$\Rightarrow \; \ddot{x}(t) + \frac{k}{m} \cdot x(t) = \frac{F(t)}{m} \tag{5.23}$$

For these kinds of systems, the solution can be split into 2 parts known as the *homogeneous* and the *particular solutions*. The homogeneous solution corresponds to the response of the system with frequency equal to the natural frequency of the system, whereas the particular solution has frequency equal to the driving frequency. For the driving force $F(t) = f\cos(\omega t + \theta)$, we can write the general solution as

$$x(t) = a_1 \cos(\omega_0 t + \phi) + a_2 \cos(\omega t + \theta) \tag{5.24}$$

where a_1, ϕ and θ are parameters to be solved using initial conditions. This solution can be easily checked to satisfy the above equation of motion. As for a_2, it can be expressed in terms of the given parameters in the form

$$a_2 = \frac{f}{m(\omega_0^2 - \omega^2)} \tag{5.25}$$

where f is the amplitude of the driving force. Note that we are using a specific sinusoidal driving force for example but in general, we can use *Fourier's theorem* to construct an arbitrary time-dependent driving force from a linear combination of sinusoidal functions.

But looking at the denominator of a_2, we see that it blows up to infinity if $\omega = \omega_0$. ω_0 is thus known as the *resonant frequency*. But due to this lack of a comprehensive mathematical description at resonance, allow us to do a little more analysis of the function to account for this. Consider the case where ω is close to ω such that $\omega = \omega_0 - \epsilon$ where $\epsilon \ll 1$. Substituting this into the ansatz we have above gives

$$x(t) = a_1 \cos(\omega_0 t + \phi) + \frac{f\cos\big((\omega_0 - \epsilon)t + \theta\big)}{m\big(\omega_0^2 - (\omega_0 - \epsilon)^2\big)} \tag{5.26}$$

The denominator in the particular solution can be Taylor expanded and approximated up to first order as follows.

$$\begin{aligned} m\big(\omega_0^2 - (\omega_0 - \epsilon)^2\big) &= m\big(\omega_0^2 - \omega_0^2(1 - \frac{\epsilon}{\omega_0})^2\big) \\ &\approx m\big(\omega_0^2 - \omega_0^2(1 - \frac{2\epsilon}{\omega_0})\big) = 2m\epsilon\omega_0 \end{aligned} \tag{5.27}$$

Employing the trigonometric identity that convert the sum of arguments to the sum of trigono-

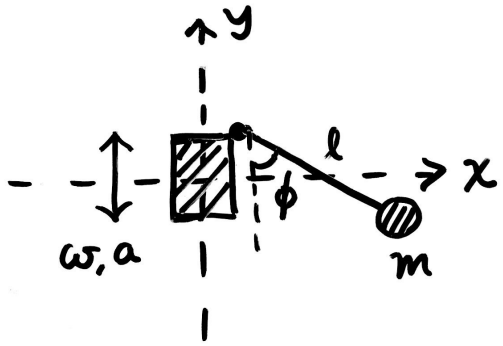

Figure 5.3: Kapitza Pendulum

metric functions, we can rewrite the ansatz as

$$x(t) = a_1' \cos(\omega_0 t + \phi') + \frac{f}{2m\epsilon\omega_0}\Big(\cos\big((\omega_0 - \epsilon)t + \theta\big) - \cos(\omega_0 t + \theta)\Big) \tag{5.28}$$

$$= a_1' \cos(\omega_0 t + \phi') + \frac{f}{2m\epsilon\omega_0}\,\mathrm{Re}\Big\{e^{i\left((\omega_0-\epsilon)t+\theta\right)} - e^{i(\omega_0 t+\theta)}\Big\} \tag{5.29}$$

$$= a_1' \cos(\omega_0 t + \phi') + \frac{f}{2m\epsilon\omega_0}\,\mathrm{Re}\Big\{e^{i(\omega_0 t+\theta)}(e^{-i\epsilon t} - 1)\Big\} \tag{5.30}$$

$$\approx a_1' \cos(\omega_0 t + \phi') + \frac{f}{2m\omega_0}\,\mathrm{Re}\Big\{-it \cdot e^{i(\omega_0 t+\theta)}\Big\} \tag{5.31}$$

Which brings us to the result

$$\boxed{x(t) = a_1' \cos(\omega_0 t + \phi') + \Big(\frac{f}{2m\omega_0}\Big) t \sin(\omega_0 t + \theta)} \tag{5.32}$$

We see here that particular solution is now scaled by a linear term in t for values of ω close to ω_0. This indeed allows our solution to have a calculable finite value at resonance.

We now look at at an example of an interesting driven system known as the *Kapitza's pendulum*.

Example

Consider a pendulum that is attached to a driven block by a pivot with a driving frequency ω and amplitude a. The pendulum is free to rotate about the pivot and has length l and mass m. Refer to figure 5.3 for a visualization of the system. The coordinates of the system are given as

$$x = l\sin(\phi) \tag{5.33}$$

$$y = -l\cos(\phi) - a\cos(\omega t) \tag{5.34}$$

$$\Rightarrow\ \dot{x} = l\sin(\phi)\dot{\phi} \tag{5.35}$$

$$\Rightarrow\ \dot{y} = l\sin(\phi)\dot{\phi} + a\omega\cos(\omega t) \tag{5.36}$$

The kinetic and potential energies of the system are then given as

$$\begin{aligned} T &= \frac{1}{2}m(\dot{x}^2 + \dot{y}^2) \\ &= \frac{1}{2}m\big(l^2\cos^2(\phi)\dot{\phi}^2 + l^2\sin^2(\phi)\dot{\phi}^2 + a^2\omega^2\sin^2(\omega t) + 2al\omega\sin(\phi)\sin(\omega t)\big) \\ &= \frac{1}{2}m\big(l^2\dot{\phi}^2 + a^2\omega^2\sin^2(\omega t) + 2al\omega\dot{\phi}\sin(\phi)\sin(\omega t)\big) \end{aligned} \tag{5.37}$$

Then since

$$\frac{d}{dt}\big(\sin(\omega t)\cos(\phi)\big) = \omega\cos(\omega t)\cos(\phi) - \sin(\omega t)\sin(\phi)\dot{\phi} \tag{5.38}$$

$$\Rightarrow\ T = \frac{1}{2}m\big(l^2\dot{\phi}^2 + a^2\omega^2\sin^2(\omega t) + 2a\omega l\cos(\omega t)\cos(\phi)\big) \tag{5.39}$$

As for the potential energy,

$$V = mgy = -mgl\cos(\phi) \tag{5.40}$$

Which gives the Lagrangian as follows.

$$\mathcal{L} = \frac{1}{2}m\big(l^2\dot{\phi}^2 + 2al\omega\cos(\phi)\sin(\omega t)\big) + mgl\cos(\phi) \tag{5.41}$$

Notice we have dropped any constant, purely time dependent and total time derivative terms as they do not affect the equations of motion. Then utilizing the Euler-Lagrange equations, we arrive

$$F_\phi = \frac{\partial\mathcal{L}}{\partial\phi} = -ma\omega l\cos(\omega t)\sin(\phi) - mgl\sin(\phi) \tag{5.42}$$

$$p_\phi = \frac{\partial\mathcal{L}}{\partial\dot{\phi}} = ml^2\dot{\phi} \quad\Rightarrow\quad \dot{p}_\phi = ml^2\ddot{\phi} \tag{5.43}$$

$$\Rightarrow\ \ddot{\phi} = -\Big(\frac{g + a\omega^2\cos(\omega t)}{l}\Big)\sin(\phi) \tag{5.44}$$

The resulting equation of motion is very similar to that of a standard pendulum system but with a time-dependent correction in the usual defined angular frequency.

For the following analysis, we are now going to consider the case where is driving frequency ω is much larger than the natural frequency of the undriven system $\omega_0 = \sqrt{g/l}$. Also, the amplitude of driving oscillations a is much smaller than the length of the pendulum l. We can now split the solution to our system into 2 parts.

$$\phi(t) = \phi_0(t) + \xi(t) \tag{5.45}$$

where ϕ_0 is concerned with the slow frequency response of the system and ξ is concerned with the fast frequency response. Looking first at the $\xi(t)$ portion of the solution, It is fair to say that the ϕ_0 portion of the solution remains relativiely 'constant' with respect to ω frequency scales. As such, we guess a solution of the form

$$\xi(t) = \frac{a}{l}\sin(\phi_0)\cos(\omega t) \tag{5.46}$$

so that a/l accounts for the amplitude, $\cos(\omega t)$ accounts for the high frequency oscillations and $\sin(\phi_0)$ accounts for the ϕ_0 dependence. We now embark on looking for an equation of motion for $\phi_0(t)$ from our guess.

$$\phi_0 = \phi - \xi \tag{5.47}$$

$$\Rightarrow \ \ddot{\phi}_0 = \ddot{\phi} - \ddot{\xi}$$

$$= -\Big(\frac{g + a\omega^2\cos(\omega t)}{l}\Big)\sin(\phi_0 + \xi) - \ddot{\xi} \tag{5.48}$$

From here, we utilize the fact that ϕ_0 acts as a constant term when we integrate over the high frequency oscillations and that $\xi \ll 1$. As such, we take the time average of $\ddot{\phi}_0$ over high frequency periods and take the Taylor approximation up to first orders in ξ to get

$$\langle\ddot{\phi}_0\rangle \approx -\langle\Big(\frac{g + a\omega^2\cos(\omega t)}{l}\Big)\big(\sin(\phi_0) + \cos(\phi_0)\xi\big)\rangle + \langle\ddot{\xi}\rangle$$

$$= -\frac{g}{l}\sin(\phi_0) - \frac{1}{2}\Big(\frac{a\omega}{l}\Big)^2\sin(\phi_0)\cos(\phi_0) \tag{5.49}$$

We can then use numerical methods to determine the trajectory of the pendulum from this and add the $\xi(t)$ solution to get an accurate picture of the behaviour of this system. However what is interesting about this system from something we can see analytically is the points stable equilibrium. Looking at the time averaged behaviour, the equilibrium points can be found by setting

$$\frac{\partial V_{\text{eff}}}{\partial\phi_0} = -ml^2\ddot{\phi}_0 = 0 \tag{5.50}$$

$$\Rightarrow \ \Big(gl + \Big(\frac{a\omega}{\sqrt{2}}\Big)^2\cos(\phi_0)\Big)\sin(\phi_0) = 0 \tag{5.51}$$

$$\Rightarrow \ \phi_0 = 0, \pi \tag{5.52}$$

We now need to check if these equilibrium points are stable by using the second derivative test. The $\phi_0 = 0$ solution is trivial and obvious so we will not check that, but it is a good reassurance that our analysis has been done right.

$$\frac{\partial^2 V_{\text{eff}}}{\partial\phi_0^2}\bigg|_{\phi_0=\pi} = gl\cos(\phi_0) + \frac{(a\omega)^2}{2}\cos(2\phi_0)\bigg|_{\phi_0=\pi} \tag{5.53}$$

$$= \frac{(a\omega)^2}{2} - gl \tag{5.54}$$

Hence we see that the $\phi_0 = \pi$ solution is stable **only** if $\boxed{(a\omega)^2 > 2gl}$. If you have not realized why this is an interesting result, what we have shown is that by vibrating our system at particular amplitudes and frequencies, we are able to stabilize an inverted pendulum! What is happening here is that the driving force has effectively modified the potential curve of the system and forced it to have an additional local minima at $\phi_0 = \pi$. We will return to this in the chapter on *tricky potentials*.

§5.2 Dissipation

We now look into systems where the energy of the system is loss to some external environment. Although dissipative systems are often not considered to be in the field of mechanics, it is essential for a comprehensive study of classical physics and thus we will look at it here. System losses could be a time dependent process and as such, we can model it as

$$T_{\text{loss}} = \int \epsilon(t) T dt \tag{5.55}$$

where ϵ is some fraction per unit time and T is the kinetic energy of the equivalent non-dissipative system. We have denoted the loss with a T because more often than not, losses occur due to some microscopic transfer of kinetic energy via heat, sounds and other forms.

As such, we can consider the Lagrangian of the larger system that includes the environment labelled as $\mathcal{L}_{\text{Dis}}$, and it would be the sum of the system Lagrangian along with the dissipated kinetic energy.

$$\begin{aligned} \mathcal{L}_{\text{Dis}} &= \mathcal{L} + \int \epsilon(t) T dt \\ \Rightarrow \quad \frac{d}{dt}\frac{\partial \mathcal{L}_{\text{Dis}}}{\partial \dot{q}} &= \frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{q}} + \frac{d}{dt}\left(\int \frac{\partial}{\partial \dot{q}} \epsilon(t) T dt \right) \end{aligned} \tag{5.56}$$

and by the Euler-Lagrange equation,

$$\begin{aligned} \frac{d}{dt}\frac{\partial \mathcal{L}_{\text{Dis}}}{\partial \dot{q}} &= \frac{\partial \mathcal{L}}{\partial q} + \frac{\partial}{\partial \dot{q}}\big(\epsilon(t) T\big) \\ \Rightarrow \quad \frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{q}} &= \frac{\partial \mathcal{L}}{\partial q} - \frac{\partial}{\partial \dot{q}}\big(\epsilon(t) T\big) \end{aligned} \tag{5.57}$$

The resulting equation (5.57) is known as the *modified Euler-Lagrange equation* and accounts for dissipation in the system of interest. Notice that it is written in terms of the Lagrangian of the system itself, **not** including the environment. In the analysis above, we have assumed a system that has one degree of freedom and only requires one dynamical variable to describe it. We can extend this to a generalized system by adding indices to equation (5.57) to get

$$\frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{q}_j} = \frac{\partial \mathcal{L}}{\partial q_j} - \frac{\partial}{\partial \dot{q}_j}\big(\epsilon(t) T\big) \tag{5.58}$$

We can also write a generalize kinetic energy term in the same way we have done in equation (2.73), replacing the matrix element coefficients notation a_{jk} with b_{jk} to avoid confusion.

$$\boxed{D = \frac{1}{2}\sum_{jk} b_{jk}\dot{q}_j\dot{q}_k} \tag{5.59}$$

$$\Rightarrow \quad \boxed{\frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{q}_j} = \frac{\partial \mathcal{L}}{\partial q_j} - \frac{\partial}{\partial \dot{q}_j}\Big(\frac{1}{2}\sum_{jk} b_{jk}\dot{q}_j\dot{q}_k\Big)} \tag{5.60}$$

where D is the *dissipation function* (energy loss from the system). Let us now look at a simple example of a dissipative system.

Example

Consider a a simple harmonic oscillator system with the given dissipation function

$$D = \frac{1}{2}b\dot{x}^2 \tag{5.61}$$

The Lagrangian for the simple harmonic oscillator is given as

$$\mathcal{L} = \frac{1}{2}m\dot{x}^2 - \frac{1}{2}kx^2 \tag{5.62}$$

$$\Rightarrow \; F_x = -kx, \quad \dot{p}_x = m\ddot{x} \tag{5.63}$$

Then substituting these into the modified Euler-Lagrange equation, we get

$$\begin{aligned} m\ddot{x} &= -kx - \frac{\partial}{\partial \dot{x}}\Big(\frac{1}{2}b\dot{x}^2\Big) \\ &= -kx - b\dot{x} \end{aligned} \tag{5.64}$$

As such, we get the equation of motion for x as

$$\boxed{\ddot{x} + 2\lambda\dot{x} + \omega_0^2 x = 0}, \quad \lambda \equiv \Big(\frac{b}{2m}\Big), \omega_0^2 \equiv \Big(\frac{k}{m}\Big) \tag{5.65}$$

which is exactly the equation of motion for a damped harmonic oscillator!

As a little refresher on damped harmonic oscillator systems, recall that there are 3 modes of damping that can occur.

1. **Under-damped Systems**: $(\lambda < \omega_0)$

   $$x(t) = ae^{-\lambda t}\cos(\omega_1 t + \phi), \quad \omega_1 \equiv \sqrt{\omega_0^2 - \lambda^2} \tag{5.66}$$

   where a is the initial amplitude of oscillation, ϕ is the phase delay found from initial conditions and ω_1 is the frequency response of the system due to both the natural and driving frequencies.

2. **Over-damped Systems**: $(\lambda > \omega_0)$

   $$x(t) = e^{-\lambda t}(a_1 e^{\beta t} + a_2 e^{-\beta t}), \quad \beta \equiv \sqrt{\lambda^2 - \omega_0^2} \tag{5.67}$$

   where a_1 abd a_2 are the amplitudes found from the initial conditions of the system.

3. **Critically Damped Systems**: $(\lambda = \omega_0)$

   $$x(t) = e^{-\lambda t}(a_1 + a_2 t) \tag{5.68}$$

   where a_1 abd a_2 are the amplitudes found from the initial conditions of the system.

§5.2.1 Energy Loss from Dissipation

In this last portion of the chapter, we will be looking at the energy loss from the system due to dissipative effects. For this, we look at the generalized form of the energy as presented in

equation (3.14).

$$\frac{d}{dt}E = \dot{q}\frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{q}} - \dot{q}\frac{\partial \mathcal{L}}{\partial q} \tag{5.69}$$

$$= \dot{q}\Big(\frac{\partial \mathcal{L}}{\partial q} - \frac{\partial}{\partial \dot{q}}D\Big) - \dot{q}\frac{\partial \mathcal{L}}{\partial q} \tag{5.70}$$

which gives the result

$$\boxed{\frac{d}{dt}E = -\dot{q}\frac{\partial}{\partial \dot{q}}D} \tag{5.71}$$

Chapter 6

Modified Potentials

In this chapter, we will be studying the effects of driving on potentials of a system. We recall that in our Lagrangian, it is possible to redefine the potential with an effective potential which reformulates our physical problem. Here, driving has an important part to play in reaching a similar reformulation of the system's potential. We begin by looking at a case study, revisiting the Kapitza pendulum (figure 5.3) that we analyzed chapter 5. Now, instead of explicitly defining the driving function, we will simply label it as $y_d(t)$ with no specific frequency or amplitude.

Chapter 7

Hamiltonian Mechanics

The final chapter in this set of notes will touch on another formalism of classical mechanics. Here, we utilize a construct known as the *Hamiltonian* to analyze the mechanics of a dynamical system. *Hamiltonian mechanics* is generally a more complicated approach and not often utilized in classical mechanics problems. Conversely, the Hamiltonian is a formalism that all of quantum mechanics is constructed on so we should not be so quick to disregards its use. We begin by building an intuition of Hamiltonians from Lagrangian mechanics which we have now familiarized ourselves with. To do this, we need to understand the *Legendre transformations*

§7.1 Legendre Transformations

Derived by Adrien-Marie Legendre, the Legendre transform is an involutive transformation on the real-valued convex functions of one real variable. Simply put, it converts one set of variables to a conjugate set of variables. In classical mechanics, we use this transformation to move between the Lagrangian and the Hamiltonian and can be thought of as similar to moving between time and frequency domains with Fourier transforms.

Definition 7.1.1. Convex Function: *Given a function $f(x)$ over a domain X, we say that $f(x)$ is convex if*

$$\forall x_1, x_2 \in X, \forall t \in [0,1] : f(tx_1 + (1-t)x_2) \leq f(tx_1) + f((1-t)x_2) \tag{7.1}$$

*$f(x)$ is **strictly convex** if the equality is dropped in (7.1) .*

Definition 7.1.2. Legendre Transform: *Given a strictly convex and differentiable function $f : \mathbb{R} \to \mathbb{R}$, the Legendre transform $G(p) = \mathbb{L}\{f(x)\}(p)$ is defined as*

$$\mathbb{L}\{f(x)\}(p) = \max_x \{xp - f(x)\} \tag{7.2}$$

$$\text{or} \quad G(p) = x_p p - f(x_p) \tag{7.3}$$

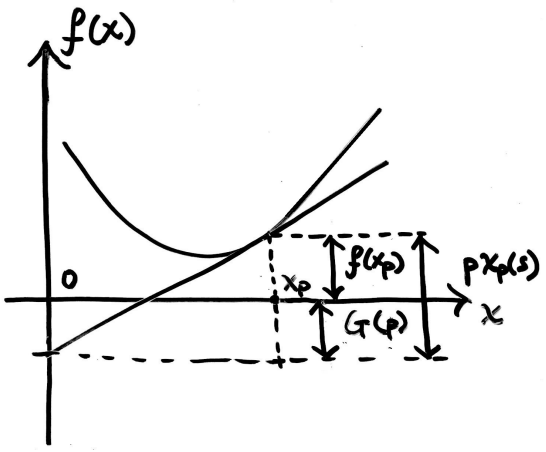

Figure 7.1: Legendre Transform Visualization

Where in (7.3), x_p solves the relation $p = f'(x_p)$. Alternative, we can rewrite (7.3) to reverse the the independent and dependent variables, which gives

$$f\big(x_p(p)\big) + G(p) = f'\big(x_p(s)\big)x_p(p) \tag{7.4}$$

This way of writing the definition allows use to use a graphical approach to understand what the Legendre transform is. As seen in figure 7.1, the Legendre transform is the difference between the height of the tangent and the function value which shows a geometric duality transformation. It is useful now to present the properties of a Legendre transform.

Properties:

1. *Preservation of Convexity*: The Legendre transform of a strictly convex function is also strictly convex.

Proof. Let $f(x)$ be a strictly convex and differentiable function, then its Legendre transform is given as

$$G(p) = x(p)p - f\big(x(p)\big) \tag{7.5}$$

Then considering the derivative of $G(p)$ with respect to p,

$$\begin{aligned}
\frac{dG(p)}{dp} &= \frac{d}{dp}\Big(x(p)p\Big) - \frac{d}{dp}f\big(x(p)\big) && (7.6)\\
&= x + x'(p)p - f'\big(x(p)\big)x'(p) && (7.7)\\
&= x + x'(p)p - x'(p)p && (7.8)\\
&= x && (7.9)
\end{aligned}$$

Now considering the second derivative,

$$\begin{aligned}
\frac{d^2G(p)}{dp^2} = x'(p) &= \frac{1}{p'(x)} && (7.10)\\
&= \frac{1}{f''(x)} > 0 && (7.11)
\end{aligned}$$

since $f(x)$ is already convex, which thus shows that G(p) also fulfills the second derivative test and is indeed strictly convex. □

Notice in the above proof that we have dropped the subscript notation from x_p for convenience. We will adopt this notation as it is inferred for its use in classical mechanics.

2. *Involution*: $G\big(G(p)\big) = \mathbb{L}\{\mathbb{L}\{f(x)\}\} = f(x)$

Proof. Let $f(x)$ be a convex differentiable function, then by definition,

$$\begin{aligned} & G(p) = \mathbb{L}\{f(x)\}(p) = x(p)f'\big(x(p)\big) - f\big(x(p)\big) \\ \Rightarrow\ & G\big(G(p)\big) = \mathbb{L}\{\mathbb{L}\{f(x)\}\} = pG'(p) - G(p) \end{aligned}$$

Then since

$$\begin{aligned} & p = f'(x), \quad G'(p) = x \\ \Rightarrow\ & G\big(G(p)\big) = xf'(x) - \Big(xf'(x) - f(x)\Big) = f(x) \end{aligned}$$

Showing that the Legendre transform is indeed an involution. □

Now that we have established the necessary formalism, we return to the physics we are trying to extract from it. As such, consider performing a Legendre transformation on the Lagrangian. As a start, we look at a 1 dimensional system for simplicity of analysis and a common form of the Lagrangian.

$$\mathcal{L}(q, \dot{q}) = \frac{1}{2}m\dot{q}^2 - v(q) \tag{7.12}$$

$$\Rightarrow\ G(p) = \mathbb{L}\{\mathcal{L}\big(q, \dot{q}(p)\big)\}(p) = p\dot{q}(p) - \mathcal{L}\big(q, \dot{q}(p)\big) \tag{7.13}$$

Where the associated variables used here mapped from those defined in the Legendre transform definition above are

$$\begin{cases} x \to \dot{q} \\ p \to p \end{cases} \tag{7.14}$$

$$\begin{aligned} \Rightarrow\ & p = \frac{\partial \mathcal{L}}{\partial \dot{q}} = m\dot{q} \\ \Rightarrow\ & \dot{q} = \frac{p}{m} \quad \Rightarrow \quad \mathcal{L}\big(q, \dot{q}(p)\big) = \frac{p^2}{2m} - V(q) \\ \Rightarrow\ & G(p) = \frac{p^2}{m} - \frac{p^2}{2m} + V(q) = \frac{p^2}{2m} + V(q) \end{aligned}$$

Then since we define that the Hamiltonian $\mathcal{H}$ is the Legendre transform of the Lagrangian,

$$\boxed{\mathcal{H} = \frac{p^2}{2m} + V(q)} \tag{7.15}$$

which is exactly $T + V$, the total energy of the system! However, the Hamiltonian is **not** always the energy of the system, like for instance when $T = T(q, \dot{q})$. Examples of this are the case when

we dealt with problems involving effective potentials. With this transformation, we have gone from a set of $\{q_j, \dot{q}_j\}$ variables to a set of $\{q_j, p_j\}$ variables. This new set of variables are known as *conjugate variables*.

So what is the Hamiltonian? It is in fact equivalent the **generalized energy** we have constructed from generalized coordinates in equation (3.14) written in conjugate variables! As such, the general Hamiltonian for an N dimensional system can be written as

$$\boxed{H(\vec{p}, \vec{q}) = \sum_j p_j \dot{q}_j(\vec{p}) - \mathcal{L}\big(\vec{q}, \dot{\vec{q}}(\vec{p})\big)} \tag{7.16}$$

§7.2 Hamilton's Equations

The next question is how do we retrieve the dynamics of a system from the Hamiltonian? The answer lies in the *Hamilton's equations* which are the Hamiltonian equivalent of the Euler-Lagrange equations. How we arrive at these are as follows. Consider the differential of a Lagrangian,

$$\begin{aligned} d\mathcal{L} &= \sum_j \frac{\partial \mathcal{L}}{\partial q_j} dq_j + \sum_j \frac{\partial \mathcal{L}}{\partial \dot{q}_j} d\dot{q}_j \\ &= \sum_j \dot{p}_j dq_j + \sum_j p_j d\dot{q}_j \end{aligned} \tag{7.17}$$

where we used the Euler-Lagrange equation to arrive at the result above. Then using the product rule identity,

$$\sum_j p_j d\dot{q}_j = d\bigg(\sum_j p_j \dot{q}_j\bigg) - \sum_j \dot{q}_j dp_j \tag{7.18}$$

$$\Rightarrow \; d\bigg(\sum_j p_j \dot{q}_j - \mathcal{L}\bigg) = d\mathcal{H} = -\sum_j \dot{p}_j dq_j + \sum_j \dot{q}_j dp_j$$

$$\Rightarrow \; \boxed{\dot{q}_j = \frac{\partial \mathcal{H}}{\partial p_j}, \quad \dot{p}_j = -\frac{\partial \mathcal{H}}{\partial q_j}} \tag{7.19}$$

Where the equations in (7.19) are the Hamilton's equations or *canonical equations*. We then use these to derive the equations of motion of a system just as the Euler-Lagrange equations would. Notice that these are first order differential equations and are usually much easier to solve analytically, but the complexity of Hamiltonian mechanics arises in the formulation of the Hamiltonian itself. Hamiltonians are usually difficult to just write down and are most easily found from the Lagrangian. With this in mind, what is useful about using the Hamiltonian rather than Lagrangian formalism is that if we have a Hamiltonian that is not an explicit function of time, it is **always** conserved! This holds even for system where the energy is not.

Theorem 7.2.1. *Given a system such that its Hamiltonian is not an explicit function of time, then it follows that the Hamiltonian is conserved.*

Proof. Looking at the definition of the total time derivative of the Hamiltonian,

$$\frac{d\mathcal{H}}{dt} = \frac{\partial \mathcal{H}}{\partial t} + \sum_j \dot{p}_j \frac{\partial \mathcal{H}}{\partial p_j} + \sum_j \dot{q}_j \frac{\partial \mathcal{H}}{\partial q_j} \tag{7.20}$$

$$= \frac{\partial \mathcal{H}}{\partial t} + \sum_j p_j \dot{q}_j - \sum_j \dot{q}_j p_j = \frac{\partial \mathcal{H}}{\partial t} \tag{7.21}$$

Hence we see that the total time derivative of the Hamiltonian is exactly equal to its partial derivative, making both terms vanish of the Hamiltonian has no explicit time dependence. This implies conservation. □

Let us now work through a simple problem and show that we can also arrive at the Lagrangian from the Hamiltonian.

Example

Recall Kepler's problem of a planet orbiting the Sun. Here, we assert that this is a non-dissipative system (energy is conserved) and that the potential due to the sun can be treated as a central potential (angular momentum is conserved). Seeing that the total energy of the system will not change in time, we know that it will not be an explicit function of time and hence we can write the Hamiltonian as

$$\mathcal{H} = \frac{p_r^2}{2\mu} + \frac{p_\phi^2}{2r^2\mu} + V(r) \tag{7.22}$$

From before, we also know that the Lagrangian of this system is given as

$$\mathcal{L} = T - V = \frac{1}{2}\mu\dot{r}^2 + \frac{1}{2}\mu(r\dot{\phi})^2 - V(r) \tag{7.23}$$

We now show that the Legendre transform of the Hamiltonian indeed returns this result. Using the canonical equations in (7.19), we get

$$\dot{r} = \frac{\partial \mathcal{H}}{\partial p_r} = \frac{p_r}{\mu}, \quad \dot{\phi} = \frac{\partial \mathcal{H}}{\partial p_\phi} = \frac{p_\phi}{r^2\mu} \tag{7.24}$$

$$\Rightarrow \quad \mathcal{H} = \frac{1}{2}\mu\dot{r}^2 + \frac{1}{2}\mu(r\dot{\phi})^2 + V(r) \tag{7.25}$$

From here, using the fact that a Legendre transform is an involution,

$$\mathbb{L}\{\mathcal{H}\} = \mathbb{L}\{\mathbb{L}\{\mathcal{L}\}\} = \mathcal{L} = (\mu\dot{r})\dot{r} + (r^2\dot{\phi}\mu)\dot{\phi} - \mathcal{H} \tag{7.26}$$

$$= (\mu\dot{r})\dot{r} + (r^2\dot{\phi}\mu)\dot{\phi} - \left(\frac{1}{2}\mu\dot{r}^2 + \frac{1}{2}\mu(r\dot{\phi})^2 + V(r)\right)$$

$$= \frac{1}{2}\mu\dot{r}^2 + \frac{1}{2}\mu(r\dot{\phi})^2 - V(r)$$

$$= T - V \tag{7.27}$$

which indeed returns the result we expected.

§7.3 Poisson Brackets

In Lagrangian and Newtonian mechanics, we saw conservation laws arise from differentiable symmetries in the system which directly follows from Noether's theorem. In this section, we look at another means to derive conserved quantities using a formalism known as *Poisson brackets*.

To understand where the Poisson bracket comes from, we first consider a generic function $f(p,q,t)$ of conjugate variables in 1 dimension. Taking its total time derivative gives

$$\begin{aligned}\frac{d}{dt}f(p,q,t) &= \frac{\partial f}{\partial t} + \dot{q}\frac{\partial f}{\partial q} + \dot{p}\frac{\partial f}{\partial p} \\ &= \frac{\partial f}{\partial t} + \frac{\partial \mathcal{H}}{\partial p}\frac{\partial f}{\partial q} - \frac{\partial \mathcal{H}}{\partial q}\frac{\partial f}{\partial p} \\ \Rightarrow \quad \frac{d}{dt}f - \frac{\partial}{\partial t}f &= \frac{\partial \mathcal{H}}{\partial p}\frac{\partial f}{\partial q} - \frac{\partial \mathcal{H}}{\partial q}\frac{\partial f}{\partial p} \equiv [\mathcal{H}, f] \end{aligned} \tag{7.28}$$

where $[\mathcal{H}, f]$ is known as the Poisson bracket of the Hamiltonian $\mathcal{H}$ with f. From this, we see that if the function f we are working with is **not** an explicit function of time, (7.28) reduces to

$$\boxed{\frac{d}{dt}f(p,q) = [\mathcal{H}, f]} \tag{7.29}$$

which implies that $f(p,q)$ is conserved if and only if $[\mathcal{H}, f] = 0$.

There is however, a case where f is a time-explicit function while having its Poisson bracket with the Hamiltonian vanish. This occurs if f is **only** time dependent, which is to say $f = f(t)$.

$$\frac{d}{dt}f(t) = \frac{\partial}{\partial t}f(t) \tag{7.30}$$

$$\Rightarrow \quad [\mathcal{H}, f] = \frac{d}{dt}f(t) - \frac{\partial}{\partial t}f(t) = 0 \tag{7.31}$$

But in this scenario, this does **not** imply that f is conserved. Equation (7.28) actually gave a specific example of a Poisson bracket. We can generalize the definition as follows.

Definition 7.3.1. Poisson Bracket: *Given 2 functions of the conjugate variables $f(\vec{p}, \vec{q}, t)$ and $g(\vec{p}, \vec{q}, t)$, their Poisson bracket is defined as*

$$[f,g]_{(p,q)} \equiv \sum_j \left(\frac{\partial f}{\partial p_j}\frac{\partial g}{\partial q_j} - \frac{\partial f}{\partial q_j}\frac{\partial g}{\partial p_j} \right) \tag{7.32}$$

The subscripts on the bracket are often dropped when the variables in use have been stated. The formalism of Poisson brackets may not seem extremely useful within the confines of the content covered in these notes, but its mathematical structures extend far into many other areas of physics (e.g. manifesting as *commutators* in quantum mechanics). It would be good to now look at several properties of Poisson brackets.

Properties:

1. *Anti-symmetric*: $[f,g] = -[g,f]$

2. *Distributive*: $[f+g,h] = [f,h] + [g,h]$
3. *Product Rule*: $[f \cdot g, h] = f[g,h] + g[f,h]$
   $\Rightarrow [f^2, g] = 2f[f,g]$
4. *Jacobi's Identity*: $[f,[g,h]] + [g,[h,f]] + [h,[f,g]] = 0$
5. *Canonical Relations*: $[q_j, q_k] = [p_j, p_k] = 0, \quad [p_j, q_k] = \delta_{jk}$
6. $[f, \alpha] = 0$, where α is some constant.
7. $[f,f] = 0$
8. $[f, q_j] = \frac{\partial f}{\partial p_j}, \quad [f, p_j] = -\frac{\partial f}{\partial q_j}$

The proof for these properties are generally considered trivial (albeit some being rather tedious) and do not reveal anything much about the physics of things, hence will not be shown here. From these properties (especially the Jacobi identity), we can construct the *Poisson's theorem* which gives us an invaluable tool to derive conserved quantities given others.

Theorem 7.3.1. *Given 2 integrals of motion f and g that are not explicit functions of time i.e.*

$$\frac{df}{dt} = \frac{\partial f}{\partial t} = 0, \quad \frac{dg}{dt} = \frac{\partial g}{\partial t} = 0 \tag{7.33}$$

then it follows that the Poisson bracket of f with g is also conserved.

$$\Rightarrow \frac{d}{dt}[f,g] = 0 \tag{7.34}$$

Proof. By the properties of f and g asserted at the beginning of the theorem, it follows that

$$[\mathcal{H}, f] = [\mathcal{H}, g] = 0 \tag{7.35}$$

Then utilizing the Jacobi identity, we get

$$[\mathcal{H},[f,g]] + [f,[g,\mathcal{H}]] + [g,[\mathcal{H},f]] = 0 \tag{7.36}$$
$$\Rightarrow \; [\mathcal{H},[f,g]] = 0 \tag{7.37}$$
$$\Rightarrow \; \frac{d}{dt}[f,g] = 0 \tag{7.38}$$

where we used the fact that f and g are not explicit functions of time. □

Let us now look at applications of Poisson's theorem by working through several simple example.

Example 1

Suppose we have a system whereby we know the x and y components of angular momenta are conserved. As a recap, the Cartesian coordinate form of angular momentum is defined

as

$$\vec{L} = \vec{r} \times \vec{p} = \begin{bmatrix} yp_z - zp_y \\ zp_x - xp_z \\ xp_y - yp_x \end{bmatrix} \tag{7.39}$$

Since L_x and L_y are not explicit functions of time and are taken to be conserved, then it follows from Poisson's theorem that $[L_x, L_y]$ is also conserved. We now see what this additional conserved quantity is.

$$\frac{\partial L_x}{\partial \vec{r}} = \begin{bmatrix} \partial L_x/\partial x \\ \partial L_x/\partial y \\ \partial L_x/\partial z \end{bmatrix} = \begin{bmatrix} 0 \\ p_z \\ -p_y \end{bmatrix}, \quad \frac{\partial L_y}{\partial \vec{r}} = \begin{bmatrix} \partial L_y/\partial x \\ \partial L_y/\partial y \\ \partial L_y/\partial z \end{bmatrix} = \begin{bmatrix} -p_z \\ 0 \\ p_x \end{bmatrix} \tag{7.40}$$

$$\frac{\partial L_x}{\partial \vec{p}} = \begin{bmatrix} \partial L_x/\partial p_x \\ \partial L_x/\partial p_y \\ \partial L_x/\partial p_z \end{bmatrix} = \begin{bmatrix} 0 \\ -z \\ y \end{bmatrix}, \quad \frac{\partial L_y}{\partial \vec{p}} = \begin{bmatrix} \partial L_y/\partial p_x \\ \partial L_y/\partial p_y \\ \partial L_y/\partial p_z \end{bmatrix} = \begin{bmatrix} z \\ 0 \\ x \end{bmatrix} \tag{7.41}$$

$$\Rightarrow \; [L_x, L_y] = \frac{\partial L_x}{\partial \vec{p}} \cdot \frac{\partial L_y}{\partial \vec{r}} - \frac{\partial L_x}{\partial \vec{r}} \cdot \frac{\partial L_y}{\partial \vec{p}} = yp_x - xp_y = -L_z \tag{7.42}$$

$$\Rightarrow \; \frac{d}{dt}[L_x, L_y] = -\frac{d}{dt} L_z = 0 \tag{7.43}$$

Hence we have shown that L_z is also conserved along with L_x and L_y! Note that the initial assertion of L_x and L_y being conserved is not the same as them being constrained. Constraining L_x and L_y may not lead to the conservation of L_z.

Example 2

Consider the 1 dimensional simple harmonic oscillator system. The Hamiltonian is given as

$$\mathcal{H} = \frac{p^2}{2m} + \frac{1}{2} m\omega^2 x^2 \tag{7.44}$$

We can perform some manipulations on the Hamiltonian to get

$$\begin{aligned} \mathcal{H} &= \frac{1}{2m}\big(p^2 + (m\omega x)^2\big) \\ &= \frac{1}{2m}(p + im\omega x)(p - im\omega x) \end{aligned} \tag{7.45}$$

The 2 conjugate quantites $p \pm im\omega x$ are not very physical in classical mechanics but are important in quantum mechanics as they form operators known as the *ladder operators*. In any case, let us consider some function of $p + im\omega x$ and time,

$$u(p, q, t) = f(p + im\omega x) + g(t) \tag{7.46}$$

Then the Poisson bracket of the Hamiltonian and u gives

$$\begin{aligned} [\mathcal{H}, u] &= (m\omega^2 x) f'(p + im\omega x) - \big(\frac{p}{m}\big)\big(im\omega f'(p + im\omega x)\big) \\ &= -i\omega(p + im\omega x) f'(p + im\omega x) \end{aligned} \tag{7.47}$$

For convenience, we want to choose the functional form of $f(p + im\omega x)$ such that

$$(p + im\omega x) f'(p + im\omega x) = 1 \tag{7.48}$$

$$\Rightarrow \; u(p, q, t) = \ln(p + im\omega x) - i\omega t \tag{7.49}$$

where we have chosen $g(t) = -i\omega t$ (we will see why soon). The Poisson bracket of $\mathcal{H}$ with $u(p + im\omega x, t)$ is now

$$\begin{aligned}[\mathcal{H}, u] &= \left(\frac{p}{m}\right)\left(\frac{im\omega}{p + im\omega x}\right) - \frac{m\omega^2 x}{p + im\omega x} \\ &= \frac{i\omega p - m\omega^2 x}{p + im\omega x} = i\omega \end{aligned} \tag{7.50}$$

and the partial time derivative of $u(p + im\omega x, t)$ is

$$\frac{\partial}{\partial t} u(p + im\omega x) = -i\omega \tag{7.51}$$

Then from (7.28), we substitute our results to get

$$\begin{aligned}\frac{d}{dt} u(p + im\omega x) &= [\mathcal{H}, u] + \frac{\partial}{\partial t} u(p + im\omega x) \\ &= i\omega - i\omega = 0 \end{aligned} \tag{7.52}$$

Hence, this means that this strange quantity $\boxed{u(p + im\omega x) = \ln(p + im\omega x) - i\omega t}$ is in fact a **conserved** quantity in 1-D simple harmonic oscillator systems.

Example 3

Consider again a central potential, now in the context of a planetary orbit around the sun (Kepler's problem). The Hamiltonian for such a system is given as

$$\mathcal{H} = \frac{p^2}{2m} - \frac{k}{r} = \left(\frac{p_x^2 + p_y^2 + p_z^2}{2m}\right) - \left(\frac{k}{\sqrt{x^2 + y^2 + z^2}}\right) \tag{7.53}$$

We will be working with Cartesian coordinates in this example as it will inherently simplify our work. Recall also the Laplace-Runge-Lenz vector we defined in (4.55). We will show again that it is indeed a conserved quantity using equation (7.29) and knowing that the Hamiltonian for this system is conserved. Choosing our Cartesian coordinates such that $\vec{L}$ points in the $\hat{z}$-direction, we can thus write

$$\vec{L} = L_z \hat{z}, \quad \left|\vec{L}\right| \equiv l \tag{7.54}$$

This means that our orbit lies in the x, y-plane and $z = p_z = 0$. This therefore reduces our Hamiltonian and the Laplace-Runge-Lenz vector to

$$\mathcal{H} = \frac{p_x^2 + p_y^2}{2m} - \frac{k}{\sqrt{x^2 + y^2}} \tag{7.55}$$

$$\begin{aligned}\vec{A} &= \begin{bmatrix} p_x \\ p_y \\ 0 \end{bmatrix} \times \begin{bmatrix} 0 \\ 0 \\ L_z \end{bmatrix} - \frac{mk}{\sqrt{x^2 + y^2}} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \\ &= \begin{bmatrix} L_z p_y - mkx/\sqrt{x^2 + y^2} \\ -l_z p_x - mky/\sqrt{x^2 + y^2} \\ 0 \end{bmatrix} \end{aligned} \tag{7.56}$$

Notice an elegant symmetry in the 2 quantities above between the x and y associated variables, where the mapping of $x \to y$ and $y \to x$ gives back the equivalent equations. As such, showing that the Poisson bracket of the Hamiltonian with one of the components of $\vec{A}$ vanishes is sufficient to prove that all components of $\vec{A}$ are conserved. We will pick A_x for our analysis. We now compute the Poisson bracket as follows,

$$\frac{\partial \mathcal{H}}{\partial x} = \frac{kx}{(x^2+y^2)^{\frac{3}{2}}}, \quad \frac{\partial \mathcal{H}}{\partial p_x} = \frac{p_x}{m} \tag{7.57}$$

$$\frac{\partial \mathcal{H}}{\partial y} = \frac{ky}{(x^2+y^2)^{\frac{3}{2}}}, \quad \frac{\partial \mathcal{H}}{\partial p_y} = \frac{p_y}{m} \tag{7.58}$$

$$\frac{\partial A_x}{\partial x} = -\frac{mk}{\sqrt{x^2+y^2}} + \frac{mkx^2}{(x^2+y^2)^{\frac{3}{2}}}, \quad \frac{\partial A_x}{\partial p_x} = 0 \tag{7.59}$$

$$\frac{\partial A_x}{\partial y} = \frac{mkxy}{(x^2+y^2)^{\frac{3}{2}}}, \quad \frac{\partial A_x}{\partial p_y} = L_z \tag{7.60}$$

Then plugging these into the Poisson bracket formula, we get

$$\begin{aligned}
[\mathcal{H}, A_x] &= \left(\frac{\partial \mathcal{H}}{\partial p_x}\frac{\partial A_x}{\partial x} - \frac{\partial \mathcal{H}}{\partial x}\frac{\partial A_x}{\partial p_x}\right) + \left(\frac{\partial \mathcal{H}}{\partial p_y}\frac{\partial A_x}{\partial y} - \frac{\partial \mathcal{H}}{\partial y}\frac{\partial A_x}{\partial p_y}\right) \\
&= \left(\frac{p_x}{m}\right)\left(-\frac{mk}{\sqrt{x^2+y^2}} + \frac{mkx^2}{(x^2+y^2)^{\frac{3}{2}}}\right) \\
&+ \left(\frac{p_y}{m}\right)\left(\frac{mkxy}{(x^2+y^2)^{\frac{3}{2}}}\right) - \left(\frac{ky}{\sqrt{x^2+y^2}}\right)(L_z) \\
&= \left(\frac{p_y kxy + p_x kx^2}{(x^2+y^2)^{\frac{3}{2}}}\right) - \left(\frac{p_x k + kyL_z}{\sqrt{x^2+y^2}}\right) \\
&= -\left(\frac{kyL_z}{\sqrt{x^2+y^2}}\right) + \left(\frac{p_y kxy + p_x kx^2 - p_x k(x^2+y^2)}{(x^2+y^2)^{\frac{3}{2}}}\right) \\
&= -\left(\frac{kyL_z}{\sqrt{x^2+y^2}}\right) + \left(\frac{kyL_z}{\sqrt{x^2+y^2}}\right) = 0
\end{aligned} \tag{7.61}$$

Proving that A_x is indeed conserved and thus, so is $\vec{A}$.

It is now good to establish, just as we have to Newtonian and Lagrangian mechanics, a systematic **work-flow** for Hamiltonian mechanics problems. The sequence of steps are as follows.

1. Write down the kinetic energy T and potential energy V of the system in Cartesian coordinates.

2. Write the coordinate transformations to generalized coordinates $\{q_j, \dot{q}_j\}$ best suited for the problem.

3. Rewrite the kinetic energy, potential energy and Lagrangian expressed in terms of the chosen generalized coordinates $T(\vec{q}, \dot{\vec{q}})$, $V(\vec{q})$ and $\mathcal{L}(\vec{q}, \dot{\vec{q}})$.

4. Compute the conjugate momenta $p_j = \frac{\partial \mathcal{L}}{\partial \dot{q}_j}$ of the system.

5. Re-express the kinetic energy in terms of the generalized coordinates and conjugate momenta $T(\vec{p}, \vec{q})$.

6. Formulate the Hamiltonian $\mathcal{H}(\vec{p}, \vec{q}) = T(\vec{p}, \vec{q}) + V(\vec{q})$ and utilize the Hamilton's equations of motion (7.19) to construct the equations of motion.

Looking through the steps above, it seems rather redundant to utilize the Hamiltonian approach if we first have to formulate the Lagrangian anyway. Is there then a means to construct the Hamiltonian without the Lagrangian? Turns out the answer is yes! These are primarily done in **2 methods**.

1. "*Guess and Check*"

   If the mechanical system we are dealing with is simple enough, it is often possible to determine its generalized coordinates and momenta from simply eyeballing it. Of course we cannot just stop there and blindly work with these guesses, we need a way to verify them!

   The way to go about doing this is by utilizing **Poisson brackets** properties. That is, any canonical set of variables chosen for a problem **must** satisfy the canonical relations. Presenting this explicitly, consider first having a set of canonical variables $\{q_j, p_j\}$ that we want to transform to another set of canonical variables $\{Q_j, P_j\}$. We assume that the original set $\{q_j, p_j\}$ satisfies the canonical relations. We then also require our new set $\{Q_j, P_j\}$ to do the same which means,

   $$[Q_j, Q_k]_{p,q} = 0, \quad [P_j, P_k]_{p,q} = 0, \quad [P_j, Q_k]_{p,q} = \delta_{j,k} \tag{7.62}$$

   where we have performed these Poisson brackets with respect to the old set of canonical variables $\{q_j, p_j\}$.

2. *Generator Functions*

   This next approach is far more mathematically rigorous and consistent than the previous one, making it the more widely used method of Hamiltonian formulation.

§7.4 Generator Functions

The motivation behind this goes back to a property of the Lagrangian we had raised earlier. This property say that adding a total time derivative term to a Lagrangian does **not** change the physics and resultant equations of motions. How then can we use this to our advantage? Going back to the principle of least action, we know that the variation of the action must vanish on the trajectory that minimizes the action.

$$\Rightarrow \delta \int_{t_1}^{t_2} \mathcal{L} dt = \delta \int_{t_1}^{t_2} \Big(\sum_j p_j \dot{q}_j - \mathcal{H}(\vec{p}, \vec{q}, t) \Big) dt = 0 \tag{7.63}$$

Also say that we want to go from one set of canonical variables to another, $(\vec{p}, \vec{q}) \to (\vec{P}, \vec{Q})$, which would cause our Hamiltonian to transform as $\mathcal{H}(\vec{p}, \vec{q}, t) \to \mathcal{K}(\vec{P}, \vec{Q}, t)$. As such, our variational principle can be rewritten as

$$\delta \int_{t_1}^{t_2} \Big(\sum_j P_j \dot{Q}_j - \mathcal{K}(\vec{P}, \vec{Q}, t) + \frac{d}{dt} F(\vec{q}, \vec{p}, \vec{Q}, \vec{P}, t) \Big) dt = 0 \tag{7.64}$$

where we have exploited the Lagrangian property mentioned above for some arbitrary function $F(\vec{q}, \vec{p}, \vec{Q}, \vec{P}, t)$, which is known as a *generator function*. Let us first consider a particular case where $F = F(\vec{q}, \vec{Q}, t)$.

$$dF(\vec{q}, \vec{Q}, t) = \frac{\partial F}{\partial t} + \sum_j \frac{\partial F}{\partial q_j} dq_j + \sum_j \frac{\partial F}{\partial Q_j} dQ_j \tag{7.65}$$

$$\Rightarrow \frac{d}{dt} F(\vec{q}, \vec{Q}, t) = \frac{\partial F}{\partial t} + \sum_j \dot{q}_j \frac{\partial F}{\partial q_j} + \sum_j \dot{Q}_j \frac{\partial F}{\partial Q_j} \tag{7.66}$$

$$\begin{aligned} \Rightarrow \sum_j p_j \dot{q}_j - \mathcal{H} &= \sum_j P_j \dot{Q}_j - \mathcal{K} + \left(\frac{\partial F}{\partial t} + \sum_j \dot{q}_j \frac{\partial F}{\partial q_j} + \sum_j \dot{Q}_j \frac{\partial F}{\partial Q_j} \right) \\ &= \Big(\sum_j P_j - \frac{\partial F}{\partial Q_j} \Big) \dot{Q}_j + \sum_j \frac{\partial F}{\partial q_j} \dot{q}_j - \Big(\mathcal{K} - \frac{\partial F}{\partial t} \Big) \end{aligned} \tag{7.67}$$

If we compare the right and left-hand sides of equation (7.67), we see that we get definitive relations between the canonical coordinates. These relations are

$$\sum_j \Big(P_j - \frac{\partial F}{\partial Q_j} \Big) = 0, \quad p_j = \frac{\partial F}{\partial q_j}, \quad \mathcal{H} = \mathcal{K} - \frac{\partial F}{\partial t} \tag{7.68}$$

This shows that choosing some arbitrary function $F(\vec{q}, \vec{Q}, t)$ allows use to 'generate' relations between the canonical variables, which would then allow us to construct our Hamiltonian! These coordinate transformations are known as *point transformations*. In fact, there are 4 different forms of generator functions with different associated point transformations. The one we have just done is known as a *generator function of the* 1^{st} *kind*. We will list all of them and their point transformations below.

1. *Generator Functions of the* 1^{st} *Kind*:

$$F_1 = F(\vec{q}, \vec{Q}, t) \tag{7.69}$$

$$\Rightarrow P_j = \frac{\partial F_1}{\partial Q_j}, \quad p_j = \frac{\partial F_1}{\partial q_j}, \quad \mathcal{K} = \mathcal{H} + \frac{\partial F_1}{\partial t} \tag{7.70}$$

2. *Generator Functions of the 2^{nd} Kind*:

$$F_2 = F(\vec{q}, \vec{P}, t) - \vec{Q} \cdot \vec{P} \tag{7.71}$$

$$\Rightarrow \; Q_j = \frac{\partial F_2}{\partial P_j}, \;\; p_j = \frac{\partial F_2}{\partial q_j}, \;\; \mathcal{H} = \mathcal{K} + \frac{\partial F_2}{\partial t} \tag{7.72}$$

3. *Generator Functions of the 3^{rd} Kind*:

$$F_3 = F(\vec{P}, \vec{Q}, t) + \vec{q} \cdot \vec{p} \tag{7.73}$$

$$\Rightarrow \; q_j = -\frac{\partial F_3}{\partial P_j}, \;\; P_j = -\frac{\partial F_3}{\partial Q_j}, \;\; \mathcal{H} = \mathcal{K} + \frac{\partial F_3}{\partial t} \tag{7.74}$$

4. *Generator Functions of the 4^{th} Kind*:

$$F_4 = F(\vec{p}, \vec{P}, t) - \vec{Q} \cdot \vec{P} + \vec{q} \cdot \vec{p} \tag{7.75}$$

$$\Rightarrow \; q_j = -\frac{\partial F_4}{\partial p_j}, \;\; Q_j = \frac{\partial F_4}{\partial P_j}, \;\; \mathcal{H} = \mathcal{K} + \frac{\partial F_4}{\partial t} \tag{7.76}$$

Equipped with these generator functions, the difficulty now comes in choosing which of them would be best to use for the particular system we are dealing with. In addition, the form of the generator function would have to be inferred from the transformation of canonical variables we want to achieve.

§7.5 Hamilton-Jacobi Equation

Despite appearances, the very fact that we have the freedom to choose your generator functions actually make them a powerful tool! To see this, consider a 1 degree of freedom system, having the canonical variables $\{p, q\}$. Then let's say we want to perform some point transformation on these variables to another set of canonical variables such that

$$\{p, q\} \to \{P, Q\} \tag{7.77}$$

$$\mathcal{H}(p, q, t) \to \mathcal{K}(P, Q, t) \tag{7.78}$$

Then using the generator function relation of the Legendre transforms

$$p\dot{q} - \mathcal{H}(p, q, t) = P\dot{Q} - \mathcal{K}(P, Q, t) + \frac{d}{dt}F(p, q, P, Q, t) \tag{7.79}$$

If we pick specifically a generator function of the 2^{nd} kind, this would give the point transforms,

$$F_2 = F(q, P, t) - QP \tag{7.80}$$

$$\Rightarrow \; Q = \frac{\partial F_2}{\partial P}, \;\; p = \frac{\partial F_2}{\partial q}, \;\; \mathcal{K} = \mathcal{H} + \frac{\partial}{\partial t}F \tag{7.81}$$

From here, due to our liberty to choose any generator function of our choice that includes a subset of all the canonical variables, let us choose F such that

$$\mathcal{K}(P, Q, t) = \mathcal{H}(p, q, t) + \frac{\partial}{\partial t}F(p, q, P, Q, t) = 0 \tag{7.82}$$

An immediate result of this choice guarantees that our new canonical variables $\{P, Q\}$ are **conserved quantities**, since

$$\dot{P} = -\frac{\partial \mathcal{K}}{\partial Q} = 0, \quad \dot{Q} = \frac{\partial \mathcal{K}}{\partial P} = 0 \tag{7.83}$$

keeping in mind that P and Q are constants, then we can write the generator as

$$F(q, P, t) = S(q, t) + A \tag{7.84}$$

where A is some constant term. Finally, since

$$p = \frac{\partial F_2}{\partial q} = \frac{\partial S}{\partial q}, \quad Q = \frac{\partial S}{\partial P} \tag{7.85}$$

$$\Rightarrow \boxed{\mathcal{H}(q, \frac{\partial s}{\partial q}, t) + \frac{\partial S}{\partial t} = 0} \tag{7.86}$$

where (7.86) is known as the *Hamilton-Jacobi equation*. As mentioned earlier, what is good about picking a generator that satisfies this form is that it allows us to **always** retrieve canonical variables which are conserved quantites.

It was not by random selection that we have chosen S as the name of our generator function above. To see why we have done so, consider the following.

$$\begin{aligned} & S = S(q, t) \\ \Rightarrow \quad & \frac{dS}{dt} = \dot{q}\frac{\partial S}{\partial q} + \frac{\partial S}{\partial t} = \dot{q}p - \mathcal{H}(p, q, t) \end{aligned}$$

where we used (7.85) and (7.86). Then by the definition of a Legendre transformation,

$$\begin{aligned} & \frac{dS}{dt} = \mathcal{L}(q, \dot{q}, t) \\ \Rightarrow \quad & S(q, t) = \int_{t_1}^{t_2} \mathcal{L}(q, \dot{q}, t) dt \end{aligned} \tag{7.87}$$

which is exactly the definition of the Action (2.28)! Allow us now do an example to show the usefulness of the Hamilton-Jacobi equation.

Example

Consider once again the 1-D simple harmonic oscillator, its Hamiltonian is given as

$$\mathcal{H} = \frac{1}{2m}(p^2 + m^2\omega^2 q^2) \tag{7.88}$$

We know that energy for this system is conserved since $\mathcal{H}$ is not an explicit function of time and $\mathcal{H} = E$. Then using the identity in (7.85) and the Hamilton-Jacobi equation

(7.86), we get

$$\frac{1}{2m}\Big(\Big(\frac{\partial S}{\partial q}\Big)^2 + m^2\omega^2 q^2\Big) + \frac{\partial S}{\partial t} = E + \frac{\partial}{\partial t}S(q,P,t) = 0 \tag{7.89}$$

$$\Rightarrow\ S(q,P,t) = -E\int \cdot dt = -Et + W(P,q) \tag{7.90}$$

$$\Rightarrow\ \frac{\partial}{\partial q}S(q,P,t) = \frac{\partial}{\partial q}W(P,q),\quad \boxed{P = E} \tag{7.91}$$

Substituting this result back into the Hamilton-Jacobi expression, we get

$$\begin{aligned}
&\Rightarrow\ \frac{1}{2m}\Big(\Big(\frac{\partial W}{\partial q}\Big)^2 + m^2\omega^2 q^2\Big) = E \\
&\Rightarrow\ \frac{\partial}{\partial q}W(p,q) = \sqrt{2mE - m^2\omega^2 q^2} \\
&\Rightarrow\ W(p,q) = \sqrt{2mE}\int dq\sqrt{1 - \frac{m\omega^2 q^2}{2E}}
\end{aligned} \tag{7.92}$$

Then looking at the other canonical conserved quantity, we have

$$\begin{aligned}
Q = \frac{\partial S}{\partial P} &= \frac{\partial S}{\partial E} = -t + \frac{\partial W(E,q)}{\partial E} \\
\boxed{Q = -t + \sqrt{\frac{m}{2E}}\int \frac{dq}{\sqrt{1 - \frac{m\omega^2 q^2}{2E}}}}&
\end{aligned} \tag{7.93}$$

Integrating and rearranging the equation, we arrive at

$$t + Q = \frac{1}{\omega}\sin^{-1}\Big(q\sqrt{\frac{m\omega^2}{2P}}\Big) \tag{7.94}$$

$$\Rightarrow\ q = \sqrt{\frac{2P}{m\omega^2}}\sin\big(\omega(t+Q)\big) \tag{7.95}$$

$$\Rightarrow\ p = m\dot{q} = \sqrt{2mP}\cos\big(\omega(t+Q)\big) \tag{7.96}$$

once again retrieving the sinusoidal trajectories as expected with the conserved quantities P and Q (boxed). In this context, we see that the conserved quantities nicely represent the energy and phase of the system, which we know for a simple harmonic oscillator are indeed conserved.

§7.6 Infinitesimal Time Translations

In this section, we will look at how the Hamiltonian is itself a special generator function which generates small *time translations*. The idea behind this is that if we were to use the Hamiltonian as the generator of point transformations, the new canonical variables would simply be the old canonical variables shifted by a small amount in time! To show this, we first look the generator functions of the 2$^{\text{nd}}$ kind. We can actually write them in a different way from how it was

previously presented, and we will denote this alternate form with a prime.

$$F_2' = \vec{q} \cdot \vec{P} + \epsilon G(\vec{q}, \vec{P}) \tag{7.97}$$

Its associated point transformations are then

$$p_j = \frac{\partial F_2'}{\partial q_j} = P_j + \epsilon \frac{\partial G}{\partial q} \tag{7.98}$$

$$Q_j = \frac{\partial F_2'}{\partial P_j} = q_j + \epsilon \frac{\partial G}{\partial P_j} \tag{7.99}$$

Notice that in the limit as $\epsilon \to 0$, this causes $P_j \to p_j$ and hence,

$$\lim_{\epsilon \to 0} \frac{\partial G(\vec{q}, \vec{P})}{\partial P_j} = \frac{\partial G(\vec{q}, \vec{p})}{\partial p_j} \tag{7.100}$$

Now let $\epsilon = dt$ and $G(\vec{q}, \vec{P}) = \mathcal{H}(\vec{q}, \vec{p})$ and substitute these into the point transformation above. As a result, we get

$$\boxed{P_j = p_j - dt \frac{\partial \mathcal{H}}{\partial q_j}, \quad Q_j = q_j + dt \frac{\partial \mathcal{H}}{\partial P_j}} \tag{7.101}$$

$$\Rightarrow \quad P_j = p_j + \dot{p}_j dt, \quad Q_j = q_j + \dot{q}_j dt$$

$$\Rightarrow \quad P_j(t) = p_j(t + dt), \quad Q_j(t) = q_j(t + dt) \tag{7.102}$$

Hence, we see from (7.101) and (7.102) that the Hamiltonian is indeed the *generator of time translation.*

§7.7 Moving Forward

In this last section, we will be looking at some useful concepts derived from the study of classical mechanics, but widely used in many other areas of physics. These will only be briefly touched upon and is more so to introduce these theorems and ideas for future familiarity. The first is the notion of *phase space.*

§7.7.1 Phase Space

In the analysis of a dynamical systems, we now know how to assess the degrees of freedom of a system and to represent them as generalized coordinates. We also know how to perform point transformations on these coordinates with the aid of generator functions into new sets of canonical variables. It is only natural to then ask, is there a plot we could construct with these abstract coordinates to give us visual insights into our system? The answer to this lies in *phase space* plots.

Definition 7.7.1. Phase Space: *In a mechanical system, phase space is a multidimensional space in which every state of the system and their trajectories over time are represented. Every canonical momentum and position variable is an axis in the phase diagram.*

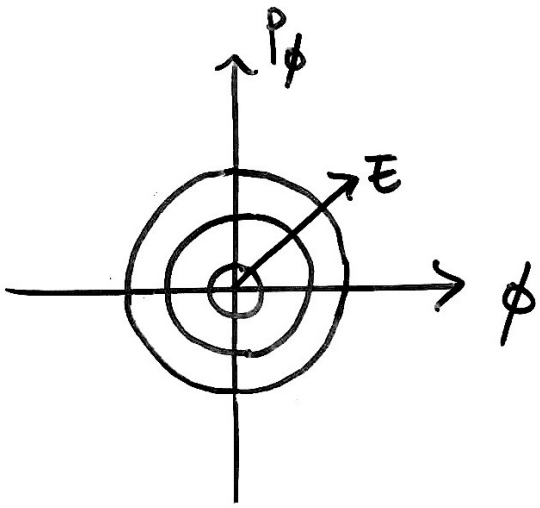

Figure 7.2: Phase Diagram of a Simple Harmonic Oscillator System

Phase diagrams are most easily read if the system can be represented by one canonical position and momentum. This is known as a *phase plane*. An example of this would be a simple pendulum system undergoing small oscillations.

Example

Consider a simple pendulum system, which gives rise to a simple harmonic potential. Taking ϕ to be the angle of deviation away from the horizontal, we get that the Hamiltonian of the system is

$$\mathcal{H} = \frac{1}{2}mR^2\dot{\phi}^2 - mgR\cos(\phi) \approx \frac{p_\phi^2}{2mR^2} - mgR \tag{7.103}$$

The corresponding plot of p_ϕ against ϕ would thus be a cyclic and circular trajectory, where the radius of 'orbit' in phase space increases with increasing energy. The phase diagram representation is given in figure 7.2.

§7.7.2 Liouville's Theorem

For this, we will simply state the theorem, its proof and the famed Liouville's equation.

Theorem 7.7.1. *Liouville's theorem says that the phase space distribution function of a system is constant throughout its trajectory.*

To understand this, we first need to know the definition of a *distribution function*.

Definition 7.7.2. *The distribution function gives the number of particles per unit volume in a single-particle phase diagram.*

From here, the proof is as such.

Proof. Consider the $2N$ dimensional vector $(\vec{p}, \vec{q})^T$. To show that the distribution of states always spans the same volume in the phase diagram, we require that the divergence of the canonical coordinate velocities vanish. Hence,

$$\begin{aligned} \boldsymbol{\nabla} \cdot (\dot{\vec{p}}, \dot{\vec{q}}) &= \sum_j \left(\frac{\partial}{\partial p_j} \dot{p}_j + \frac{\partial}{\partial q_j} \dot{q}_j \right) \\ &= \sum_j \left(- \frac{\partial}{\partial p_j} \frac{\partial \mathcal{H}}{\partial q_j} + \frac{\partial}{\partial q_j} \frac{\partial \mathcal{H}}{\partial p_j} \right) = 0 \end{aligned} \tag{7.104}$$

Indeed showing that the distribution function is constant for all time. □

This concludes our brief journey through the MIT's undergraduate second course in classical mechanics (8.223). If you have spotted any errors in these notes, please email them to *reubenw@mit.edu*. This is a **first** draft.