5210 Theoretical Mechanics

University of Colorado, Boulder (Fall 2020)

Lecturer: Prof. Victor Gurarie
Scribe: Reuben R.W. Wang

The textbook used for this course is "Classical Mechanics" by Goldstein, Poole and Safko.

All notes are taken real-time in the class (i.e. there are bound to be typos) taught by Professor Victor Gurarie. For information on this class, refer to the canvas course page. The lectures will be conducted on Zoom (MWF, 10:20am – 11:10am Mountain time) using the link https://cuboulder.zoom.us/j/92529680362.

Instructor: *Professor Victor Gurarie.*
Instructor office hours: by appointment (remote).
Instructor email: `victor.gurarie@colorado.edu`.
Personal use only. Send any corrections and comments to `reuben.wang@colorado.edu`.

Contents

Chapter 1

Analytical Mechanics

Classical mechanics is a very old subject and is the basis for many other fields of physics one would bump into. Most famously, classical mechanics is built upon Newton's 3 laws, which involve the position $\boldsymbol{x}(t)$, velocity $\dot{\boldsymbol{x}}(t)$ and acceleration $\ddot{\boldsymbol{x}}(t)$. Newton's 3 law are stated as follows:

1. *A body that is a rest or moving with constant velocity will continue to do so unless acted upon by a net external force.*
2. *The net force acting on a body is equal to the rate of change of momentum of that object.*

$$\boldsymbol{F} = \frac{d}{dt}\boldsymbol{p}. \tag{1.1}$$

   (This allows deterministic solutions to initial value problems.)

3. *If a body A, exerts a force on another, B, body B will also exert a force on body A that is equal in magnitude but opposite in direction.*

$$\boldsymbol{F}_{12} = -\boldsymbol{F}_{21}. \tag{1.2}$$

At the graduate level, it is more than likely that most students are familiar with these laws and find them easy to grasp. However, we should not be too quick to dismiss these as trivial. In fact, these laws have led to very rich, complex phenomena and even new theories of physics. For instance, the first law gave rise to a deeper question posed by Einstein, in which accelerating bodies seem to experience a force equivalent to them being in a gravitational field. This subsequently led to the equivalence principle which is the basis for Einstein's theory of gravity.

But before going beyond classical Newtonian mechanics, a full treatment of these field is only attained when all perspectives have been considered. This bring us to our first topic, analytical mechanics.

§1.1 The Principle of Least Action

In mechanics, we usually begin to describe a system by means of the constituent positions, velocities and accelerations with respect to some reference frame. However, it is in many situations more convenient if a system could be described in a frame independent way (i.e. the coordinate

frame is inconsequential to the emergent physics). In light of this, we introduce the notion of *generalized coordinates*. Generalized coordinates are a set of variables $\{q_j(t)\}$, with $j = 1, 2, 3, \ldots, N$ denoting the index to each *degree of freedom* in the physical system.

Note: The number of generalized coordinates necessary to describe a system is always equal to the number of degrees of freedom of that system.

The use of generalized coordinates are best illustrated through an example, for which we will use the problem of a double pendulum.

Example:
Consider a system of 2 pendulums connected in series, suspended from a rigid surface as illustrated in Fig. 1.1.

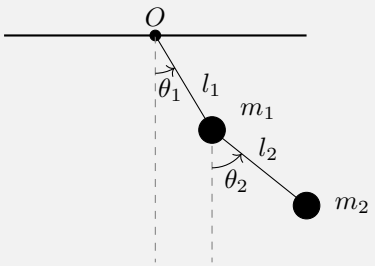

Figure 1.1: Double pendulum system.

We will take that the rods attaching to masses m_1 and m_2 of lengths l_1 and l_2 are perfectly rigid, while the motion of each pendulum is confined within a 2-dimensional plane. This implies that this dynamical system can be fully described by just 2 degrees of freedom. 2 degrees of freedom imply that only 2 generalized coordinates q_1, q_2, are required to describe this system. These 2 generalized coordinates could simply be

$$q_1 = \theta_1, \quad q_2 = \theta_2. \tag{1.3}$$

The useful thing about generalized coordinates, is that any other complete set of coordinates can also be used to described the system such as $\{x_1, x_2\}$, $\{y_1, y_2\}$, etc. However we will often find that in practice, there tends to be a most convenient set to work with.

Having the generalized coordinates of the system, the next object we will require is the *Lagrangian* (or Lagrange function), $\mathcal{L}$. The Lagrangian is a function of the generalized coordinates $q_j(t)$, as well as their time derivatives $\dot{q}_j(t)$, known as *generalized velocities*. The Lagrangian function can in general, be extremely complex and generalized to cater to many different systems. However, for the simplest case of a single particle mechanical system, the Lagrangian is simply

$$\begin{aligned} \mathcal{L}(\boldsymbol{q}, \dot{\boldsymbol{q}}) &= K - U \\ &= \frac{m\dot{\boldsymbol{q}}^2}{2} - U(\boldsymbol{q}), \end{aligned} \tag{1.4}$$

where K and U are the kinetic and potential energies of the particle respectively. In fact for much

of mechanics, this form suffices. Having the Lagrangian, we can then introduce the definition of another function known as the *action*. Before we present its definition, it would be good to understand the motivation for this.

Newton's formulation of mechanics relies on his second law to evolve a set of initial conditions in time. However, suppose that we know all the generalized coordinates of a system for some the initial and final times, t_i and t_f (i.e. we know $q_j(t_i)$ and $q_j(t_f)$ for all j). We can then come up with an alternative formulation by asking the question, how do we determine the trajectory $q_j(t)$ and interpolates the initial and final configurations of the system? To answer this, we will need a functional that considers the entire trajectory over $t \in [t_i, t_f]$, and a postulate from which we can derive an equation of motion. Well, this functional over the time trajectory will be taken as

$$S[\boldsymbol{q}(t)] = \int_{t_i}^{t_f} dt \mathcal{L}(q_j, \dot{q}_j), \tag{1.5}$$

known as the action. Note that this function has units of $\sim ET$ (energy $\cdot$ time). This is now the point that a postulate is required. The postulate is as follows:

"*Given the initial and final configurations of a mechanical system, there must be a particular set of generalized velocities that takes the initial configuration to the final one, such that the resultant trajectory minimizes the action.*"

This postulate is known as the *principle of least action*.

§1.1.1 The Euler-Lagrange Equation

How then do we find the least action? Formally, this requires a technique known as the *calculus of variations* which helps us to extremize functionals. To start, we consider a small variation $\epsilon_j(t)$, away from the optimal trajectory $q_j^{(0)}(t)$ at time t, such that

$$q_j(t) = q_j^{(0)}(t) + \epsilon_j(t). \tag{1.6}$$

Note that $\epsilon_j(t_i) = \epsilon_j(t_f) = 0$ since the end points are fixed. We now want to know how much this variation in the trajectory cause the action to vary. To do so, we substitute this variation into the action

$$S = \int_{t_i}^{t_f} \mathcal{L}\Big(q_j^{(0)}(t) + \epsilon_j(t), \dot{q}_j^{(0)}(t) + \dot{\epsilon}_j(t)\Big) dt, \tag{1.7}$$

and perform a Taylor expansion to first order which gives

$$\begin{aligned} S &\approx \int_{t_i}^{t_f} dt \left[\mathcal{L}\left(q_j^{(0)}, \dot{q}_j^{(0)}\right) + \sum_j \epsilon_j \frac{\partial \mathcal{L}}{\partial q_j^{(0)}} + \sum_j \dot{\epsilon}_j \frac{\partial \mathcal{L}}{\partial \dot{q}_j^{(0)}} \right] \\ &= \int_{t_i}^{t_f} dt \left[\mathcal{L}\left(q_j^{(0)}, \dot{q}_j^{(0)}\right) + \sum_j \epsilon_j \frac{\partial \mathcal{L}}{\partial q_j^{(0)}} - \sum_j \epsilon_j \frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{q}_j^{(0)}}\right) \right] + \epsilon_j \frac{\partial \mathcal{L}}{\partial q_j^{(0)}}\Bigg|_{t_i}^{t_f}. \end{aligned} \tag{1.8}$$

From here, we note that the variational terms vanish at the end points (t_i and t_f), which leaves the action as

$$S \approx \int_{t_i}^{t_f} dt \mathcal{L}\left(q_j^{(0)}, \dot{q}_j^{(0)}\right) + \int_{t_i}^{t_f} dt \sum_j \epsilon_j(t) \left[\frac{\partial \mathcal{L}}{\partial q_j^{(0)}} - \frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{q}_j^{(0)}}\right)\right]. \tag{1.9}$$

Extremizing this functional would then lead us to set the first-order term in the Taylor expansion to zero for all times, which leaves us with

$$\boxed{\frac{\partial \mathcal{L}}{\partial q_j} - \frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{q}_j}\right) = 0}. \tag{1.10}$$

This is the resulting equation of motion that follows from the principle of least action, and is famously known as the *Euler-Lagrange equation*. Alternatively, this could have been derived by taking the variational derivative of S with respect to $q_j(t)$ at some time t_0. Details of this alternate derivation is provided in App. A.

Note: This form of the Euler-Lagrange equations only takes into account system with conservative forces.

The similarity of the Euler-Lagrange equations to Newton's second law, also permit the definition of *canonical forces* F_j, and *canonical momenta* p_j,

$$p_j = \frac{\partial \mathcal{L}}{\partial \dot{q}_j}, \tag{1.11a}$$

$$F_j = \frac{\partial \mathcal{L}}{\partial q_j}. \tag{1.11b}$$

As a sanity check, we can see that this equation quickly reduces to Newton's second law for a single particle on a line. For a single particle in 1-dimension, we have that

$$\mathcal{L}(x, \dot{x}) = \frac{m\dot{x}^2}{2} - U(x). \tag{1.12}$$

Plugging this into the Euler-Lagrange equation then gives us

$$\begin{aligned} &\left[\frac{\partial}{\partial x}\left(\frac{m\dot{x}^2}{2} - U(x)\right)\right] - \left[\frac{d}{dt}\left(\frac{\partial}{\partial \dot{x}}\left(\frac{m\dot{x}^2}{2} - U(x)\right)\right)\right] = 0 \\ \Rightarrow \quad & m\ddot{x} = -\frac{\partial U(x)}{\partial x}, \end{aligned} \tag{1.13}$$

which is indeed Newton's second law.

§1.2 Noether's Theorem

In 1915, Emmy Noether found a deep relation between symmetries of a physical system and conserved quantities. This resulted in the famed *Neother's theorem*, which is stated below.

Theorem 1.2.1. *Every differentiable symmetry (can be parameterized by a continuous parameter) of a physical system has a corresponding conservation law.*

§1.2.1 Conservation of Momentum

The first type of symmetry we will consider will be translational symmetry. Let's suppose that we have a theory with the action

$$S_x = \int_{t_i}^{t_f} \mathcal{L}(x, \dot{x})dt, \tag{1.14}$$

having trajectory $x(t)$. Also allow us to consider that we construct another trajectory

$$y(t) = x(t) + a(t), \tag{1.15}$$

where $a(t)$ is some other trajectory. we can then also construct the action for this new trajectory

$$S_y = \int_{t_i}^{t_f} \mathcal{L}(y, \dot{y})dt = \int_{t_i}^{t_f} \mathcal{L}(x + a, \dot{x} + \dot{a})dt. \tag{1.16}$$

We say that this translation by a of the trajectory is a symmetry, if $S_x = S_y$. It is easy to see that this symmetry does not exist in systems with potential energy functions being explicit functions of the trajectory (e.g. $U(x) = mx^2/2$). However, there are systems with non-trivial potential energy functions which possess translation invariance. The most common example would be a system of particles whereby the potential is strictly a function of the distance between particles (i.e. $U = U(x_i - x_j)$ with $i \neq j$).

Noether tells us that systems with symmetries also have a conservation law, so let's see what conservation law pops out from translational symmetry. First, we assume that a is very small, which will allow us to perform analysis based on Taylor expansions. We shall then compute the difference between the actions S_x and S_y

$$\begin{aligned} S_y - S_x &= \int_{t_i}^{t_f} \mathcal{L}(y, \dot{y})dt - \int_{t_i}^{t_f} \mathcal{L}(x, \dot{x})dt \\ &= \int_{t_i}^{t_f} \mathcal{L}(x + a, \dot{x} + \dot{a})dt - \int_{t_i}^{t_f} \mathcal{L}(x, \dot{x})dt \\ &\approx \int_{t_i}^{t_f} \left[\frac{\partial \mathcal{L}}{\partial x}a(t) + \frac{\partial \mathcal{L}}{\partial \dot{x}}\dot{a}\right] dt \\ &= \int_{t_i}^{t_f} \left[\frac{\partial \mathcal{L}}{\partial x} - \frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{x}}\right] a(t)dt + \left.\frac{\partial \mathcal{L}}{\partial \dot{x}}a(t)\right|_{t_i}^{t_f}. \end{aligned} \tag{1.17}$$

Since $x(t)$ is a classical trajectory, it must satisfy the Euler-Lagrange equation, so the integrand in the expression above vanishes, leaving

$$S_y - S_x = \left.\frac{\partial \mathcal{L}}{\partial \dot{x}}a(t)\right|_{t_i}^{t_f}. \tag{1.18}$$

Since Noether's theorem speaks of a symmetry with a continuous parameter, the continuous parameter for transformation here would be a, which implies that a must not be a function of time. Then to assert a symmetry, we let $S_y - S_x = 0$, which gives us that

$$\left.\frac{\partial \mathcal{L}}{\partial \dot{x}}\right|_{t_i} = \left.\frac{\partial \mathcal{L}}{\partial \dot{x}}\right|_{t_f}. \tag{1.19}$$

Recall that $\partial\mathcal{L}/\partial\dot{x}$, is the canonical momentum, and so we have just derived the *conservation of momentum* (i.e. translational symmetry of a system implies momentum conservation)! It is not too hard to generalize this derivation to a system with multiple degrees of freedom, which leads to

$$\boxed{\left.\sum_j \frac{\partial \mathcal{L}}{\partial \dot{q}_j}\right|_{t_i} = \left.\sum_j \frac{\partial \mathcal{L}}{\partial \dot{q}_j}\right|_{t_f}}. \tag{1.20}$$

Alternatively, we could derive the conservation of momentum by virtue of what are known as *cyclic coordinates*. Cyclic coordinates are generlized coordinates that do **not** explicitly occur in the Lagrangian of the system. This means of derivation was alluded to when we mentioned that potential energy functions with explicit coordinate dependence were **not** translationally invariant.

To better understand what a cyclic coordinate implies, consider an N degree of freedom conservative system. For this, we can write the Euler-Lagrange equation as

$$\frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{q}_j} - \frac{\partial \mathcal{L}}{\partial q_j} = 0, \quad j \in \{1, ..., N\}, \tag{1.21}$$

where the Lagrangian is a function of time t, generalized coordinates q_j and generalized velocities $\dot{q}_j$. Let us now say that q_k is a cyclic coordinate for some $1 \leq k \leq N$. This means that this coordinate no longer shows up in the Lagrangian. This means that the Euler-Lagrange equation for this coordinate becomes

$$\begin{aligned} &\frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{q}_k} = 0 \\ \Rightarrow \quad &\frac{\partial \mathcal{L}}{\partial \dot{q}_k} = \text{constant}, \end{aligned} \tag{1.22}$$

once again implying the conservation of momentum.

§1.2.2 Hamiltonian Conservation

A similar derivation from the one above can be performed, but instead using time translations (i.e. the Lagrangian is not explicitly dependent on time). In this case, another conserved quantity emerges which takes the form

$$\boxed{H = \sum_j \dot{q}_j \frac{\partial \mathcal{L}}{\partial \dot{q}_j} - \mathcal{L}}. \tag{1.23}$$

This quantity is known as the *Hamiltonian* of the system and takes the role of the canonical energy. It is not to hard to see, that this function indeed reduces to the energy, $E = K + U$ for a single particle system in 1-dimension.

Note: Strictly speaking, the Hamiltonian should be written in terms of conjugate momenta variables p_j, however we will ignore this fact for now and keep the terminology adopted here.

§1.3 Constraints

In mechanical systems, objects can often only move within or along specified boundaries due to physical assets of the system. These are known as *constraints* for which they can be characterized into several types.

1. Holonomic Constraints:

   Constraints that only depend on the configuration of the system, and can thus be written in the form
   $$f(q_j) = 0; \tag{1.24}$$

2. Semi-Holonomic Constraints:

   Constraints that depend on the configuration and velocities of the system, and can thus be written in the form
   $$f(q_j, ...; \dot{q}_j) = 0; \tag{1.25}$$

3. Non-Holonomic Constraints:

   Constraints that cannot be formulated in either of the 2 forms presented above.

A holonomic constraint allows for the effective reduction in the number of degrees of freedom of the system from N to $N - 1$. These constraints can in general, be asserted by the the method of *Lagrange multipliers*. To understand this concept, we consider a simple example.

Example:

Consider a function of 2 variables $g(x_1, x_2)$ for which we want to minimize it. We know that the extrema of a function can be found using the first-order conditions
$$\frac{\partial g}{\partial x_j} = 0. \tag{1.26}$$
Suppose now that the system is also constrained by a relation between the variables x_1 and x_2 via the function
$$f(x_1, x_2) = 0, \tag{1.27}$$
which constitutes a line along the surface of the function $g(\boldsymbol{x})$. The primal problem of minimization now becomes
$$\begin{aligned} &\min_{\boldsymbol{x}} \{g(\boldsymbol{x})\} \\ \text{s.t.} \quad &f(\boldsymbol{x}) = 0. \end{aligned} \tag{1.28}$$

There are cases in which $f(\boldsymbol{x})$ is too complex for us to write any one variable in terms of the other (or others for functions of more variables). So how do we then assert this constraint in a systematic way? Well, since $\boldsymbol{\nabla} g(\boldsymbol{x})$ points in the direction of steepest ascent, the extrema along the constraint line $f(\boldsymbol{x}) = 0$, can be found when $\boldsymbol{\nabla} g(\boldsymbol{x})$ is parallel to the tangent of the constraint line, i.e.

$$\boldsymbol{\nabla} f(\boldsymbol{x}) \parallel \boldsymbol{\nabla} g(\boldsymbol{x}) \tag{1.29}$$

$$\Rightarrow \quad \frac{\partial g(\boldsymbol{x})}{\partial x_j} = \lambda \frac{\partial f(\boldsymbol{x})}{\partial x_j}, \tag{1.30}$$

where λ is known as the Lagrange multiplier. With this knowledge, we can then construct a new Lagrangian function $\tilde{\mathcal{L}}$, which incorporates the holonomic constraints

$$\tilde{\mathcal{L}} = g(\boldsymbol{x}) - \lambda f(\boldsymbol{x}). \tag{1.31}$$

The first-order condition on this function would then be

$$\frac{\partial \tilde{\mathcal{L}}}{\partial x_j} = 0, \quad \frac{\partial \tilde{\mathcal{L}}}{\partial \lambda} = 0. \tag{1.32}$$

To generalize the method above to mechanics, we can generalize to a function of many variables $\mathcal{L}(\boldsymbol{q}, \dot{\boldsymbol{q}}) = \mathcal{L}(q_1, \ldots, q_N; \dot{q}_1, \ldots, \dot{q}_N)$ and several constraint functions $f_j(\boldsymbol{q}; \dot{\boldsymbol{q}}) = 0$. The new optimization problem would now be,

$$\boxed{\tilde{\mathcal{L}}(\boldsymbol{q}, \dot{\boldsymbol{q}}) = \mathcal{L}(\boldsymbol{q}, \dot{\boldsymbol{q}}) - \sum_j \lambda_j(t) f_j(\boldsymbol{q}; \dot{\boldsymbol{q}})}, \tag{1.33}$$

where the Lagrangian multipliers are taken as time-dependent so that we can vary them in the action (i.e. λ now acts as another coordinate in the modified Lagrangian).

§1.4 Central Potentials

We are now going to consider the motion of a particle due to a *central potential*. Central potentials are potentials such that

$$U(\boldsymbol{r}) = U(r), \tag{1.34}$$

i.e. it is only a function of the magnitude of the distance from the center. More generally, we will also see that a 2 body problem in which the potential can be written as $U(\boldsymbol{r}_1, \boldsymbol{r}_2) = u(|\boldsymbol{r}_1 - \boldsymbol{r}_2|)$, can be similar reduced to a central potential problem with 1 particle. To do this effective recasting of the 2-body problem, we make the definitions

$$\boldsymbol{r} \equiv \boldsymbol{r}_1 - \boldsymbol{r}_2 \tag{1.35a}$$

$$\boldsymbol{R} \equiv \frac{m_1 \boldsymbol{r}_1 + m_2 \boldsymbol{r}_2}{m_1 + m_2}, \tag{1.35b}$$

which we call the *relative position* and *center of mass (CoM) position* respectively. The inverse transformations of these are given as

$$\boldsymbol{r}_1 = \boldsymbol{R} + \frac{m_2}{m_1+m_2}\boldsymbol{r}, \tag{1.36a}$$

$$\boldsymbol{r}_2 = \boldsymbol{R} - \frac{m_1}{m_1+m_2}\boldsymbol{r}. \tag{1.36b}$$

With this, we get that the 2-body action then becomes

$$\begin{aligned} S &= \int dt \left[\frac{m_1\dot{\boldsymbol{r}}_1^2}{2} + \frac{m_2\dot{\boldsymbol{r}}_2^2}{2} - U(|\boldsymbol{r}_1 - \boldsymbol{r}_2|)\right] \\ &= \int dt \left[\frac{M\dot{\boldsymbol{R}}^2}{2} + \frac{\mu\dot{\boldsymbol{r}}^2}{2} - U(r)\right], \end{aligned} \tag{1.37}$$

where $M = m_1+m_2$ and $\mu = m_1m_2/(m_1+m_2)$ are the total mass and *reduced mass* respectively. Sp we see that the motion of $\boldsymbol{R}$ and $\boldsymbol{r}$ completely separate, allowing us to solve the relative and CoM motion independently for which the relative motion exactly follows that of a single particle in a central potential! For the rest of the analysis, we will simply ignore the center of mass motion and focus on the central potential problem. Just looking at the Lagrangian of this system

$$\mathcal{L} = \frac{\mu\dot{\boldsymbol{r}}^2}{2} - U(r), \tag{1.38}$$

we can immediately see that the Hamiltonian is conserved (due to no explicit time-dependence) but not the momentum (due to lack of translational symmetry). There is however, another conserved quantity that arises due to the *rotational symmetry* of this system. This conserved quantity is in fact the *angular momentum*.

§1.4.1 Conservation of Angular Momentum

Consider a vector $\boldsymbol{r}$, that undergoes a rotation transformation R, such that

$$r'_j = R_{jk}r_k, \tag{1.39}$$

where quantities are summed over repeated indices. For R to be a rotation transformation, it must satisfy orthogonality, which implies $RR^T = R^TR = \mathbb{I}$. From these, it then follows that

$$\begin{aligned} &|\boldsymbol{r}'| = |\boldsymbol{r}| \\ \Rightarrow\quad &\left|\dot{\boldsymbol{r}}'\right| = |\dot{\boldsymbol{r}}|. \end{aligned} \tag{1.40}$$

To now derive the associated conserved quantity, we consider an infinitesimal rotation of the vector $\boldsymbol{r}$ by a small amount ϵ such that

$$r'_j = r_j + \epsilon\cdot\omega_{jk}r_k, \tag{1.41}$$

where ω_{jk} is a transformation matrix that satisfies antisymmetry, $\omega_{jk} = -\omega_{kj}$, and $R(\epsilon) \approx \mathbb{I}+\epsilon\omega$. The antisymmetry relation also results in

$$\omega_{jk}r_jr_k = 0, \tag{1.42}$$

since ω_{jk} is antisymmetric while $r_j r_k$ is symmetric. Now, we attempt to compute the difference between the actions $S[r'_j]$ and $S[r_j]$,

$$\begin{aligned} S[r'_j] - S[r_j] &= \int_{t_i}^{t_f} \left[\frac{1}{2}\mu \sum_j \left(\dot{r}_j + \epsilon\, \omega_{jk}\dot{r}_k + \dot{\epsilon}\omega_{jk} r_k\right)^2 - U(|r_j + \epsilon \cdot \omega_{jk} r_k|) \right] dt - S[r_j] \\ &\approx \frac{\mu}{2} \int_{t_i}^{t_f} \sum_j \dot{r}_j \left(\epsilon\, \omega_{jk}\dot{r}_k + \dot{\epsilon}\, \omega_{jk} r_k\right) dt \\ &= \frac{\mu}{2} \int_{t_i}^{t_f} \sum_j \dot{\epsilon}\, \omega_{jk}\dot{r}_j r_k \\ &= \frac{\mu}{2} \left[\epsilon\, \dot{r}_j \omega_{jk} r_k |_{t_i}^{t_f} - \int_{t_i}^{t_f} \epsilon \frac{d}{dt} \left(\dot{r}_j \omega_{jk} r_k\right) dt \right]. \end{aligned} \tag{1.43}$$

Asserting that this difference vanishes and that ϵ is in fact a constant, we finally get

$$\mu \dot{r}_j \omega_{jk} r_k = \text{constant}. \tag{1.44}$$

We can then recognize that $\mu \dot{r}_j \omega_{jk} r_k$ is in fact the angular momentum of the system, so this shows rotational invariance implies conservation of angular momentum, L. Note that we have implicitly assumed rotations about the z-axis here ($\omega_{jk} = \varepsilon_{jk3}$, where ε is the *Levi-Civita symbol*), but more generally, we have

$$\boxed{L_l = \mu \varepsilon_{ljk} \dot{r}_j r_k = [\boldsymbol{r} \times \boldsymbol{p}]_l}. \tag{1.45}$$

So going back to the central potential problem, we see that since $\mathcal{L}$ is only a function of $|\boldsymbol{r}|$ and $|\dot{\boldsymbol{r}}|$, angular momentum of the system is indeed conserved. Furthermore, this implies that the motion can be further reduced to a 2D plane that is orthogonal to the angular momentum. In light of this, it is then best to work in polar coordinates (r, ϕ), in this plane for which we have

$$\dot{\boldsymbol{r}}^2 = \dot{r}^2 + r^2 \dot{\phi}^2 \tag{1.46}$$

$$\Rightarrow \quad S = \int dt \left[\frac{\mu}{2} \left(\dot{r}^2 + r^2 \dot{\phi}^2 \right) - U(r) \right] \tag{1.47}$$

$$\Rightarrow \quad \frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \dot{\phi}} = 0$$

$$\Rightarrow \quad \boxed{\mu r^2 \dot{\phi} = \ell}, \tag{1.48}$$

where $\ell = |\boldsymbol{L}|$ is a constant of motion. This renders the Hamiltonian

$$H = \frac{\mu \dot{r}^2}{2} + U_{\text{eff}}(r); \tag{1.49}$$

$$\text{where} \quad U_{\text{eff}}(r) \equiv U(r) + \frac{\ell^2}{2\mu r^2}, \tag{1.50}$$

where $U_{\text{eff}}(r)$ is an *effective potential* which accounts for the fictitious centrifugal force. In fact, because a 2-body problem in 3D has in general 6 degrees of freedom, while the assertion of a central potential enforce 4 conserved quantities (E and $\boldsymbol{L} = [L_x, L_y, L_z]$), this results in only 2 remaining degrees of freedom which can indeed be specified by r and ϕ.

§1.4.2 Kepler's Laws of Planetary Motion

One of the earliest triumphs of physics in the scientific revolution (16th to 17th century) was the empirical evidence which led to the statement of *Kepler's 3 laws of planetary motion*, which was later derived and consistent with Newton's formulation of classical mechanics. Kepler's 3 laws model the motion of the Earth around the Sun, and are stated succinctly as follows:

1. Planetary orbits are elliptical.
2. The areas swept out per unit time around the Sun are constant ($\frac{d}{dt}A = \text{constant}$).
3. The square of a planet's orbital period around the Sun is proportional to the cube of the semi-major axis ($T^2 \propto a^3$).

In this section, we embark on an involved analysis of these laws in a constructive manner. To provide a more coherent picture, We begin by first proving Kepler's second law of planetary motion.

Kepler's Second Law: The area swept out per unit time of a planet orbiting around the Sun is constant.

Consider 2 bodies of masses M (the Sun) and m (the planet) in orbit about their collective center of mass. Then by conservation of angular momentum, we have that

$$\mu r^2 \dot{\phi} = \ell = \text{constant}. \tag{1.51}$$

From here, we note that in the plane of the trajectory, a differential segment of area dA, swept out in some differential time interval dt would be given by

$$dA = \frac{r \cdot r d\phi}{2} \tag{1.52}$$

$$\Rightarrow \quad \frac{dA}{dt} = \frac{r^2\dot{\phi}}{2} = \frac{\ell}{2\mu}. \tag{1.53}$$

This then immediately grants us that

$$\boxed{\frac{dA}{dt} = \frac{\ell}{2\mu} = \text{constant}}. \tag{1.54}$$

We now proceed to derive Kepler's first law using the result obtained from the second law.

Kepler's First Law: The orbits of planets around the Sun are elliptical.

For convenience, we will be calling the conserved quantity earlier derived as $\frac{r^2}{2}\frac{d\theta}{dt} = \frac{dA}{dt} = \frac{h}{2}$. Considering the radial acceleration and Newton's law of gravitation, we have:

$$\ddot{r} - r\dot{\theta}^2 = -\frac{G\mu}{r^2}. \tag{1.55}$$

From Kepler's second law, we have that $r^2\dot{\phi} = \ell/\mu$ and hence, $r\dot{\phi}^2 = \ell^2/(\mu^2 r^3)$. Making use of a simple identity, we get:

$$\dot{r} = -r^2 \frac{d}{dt}\frac{1}{r} = \frac{\ell}{\mu\dot{\phi}}\frac{d}{dt}\frac{1}{r} = -\frac{\ell}{\mu}\frac{d}{d\phi}\frac{1}{r} \tag{1.56}$$

Taking the another derivative of this, we get:

$$\begin{aligned} \ddot{r} &= \frac{d}{dt}\left(-h\frac{d}{d\theta}\frac{1}{r}\right) \\ &= -\frac{\ell^2}{\mu^2 r^2}\frac{d^2}{d\phi^2}\frac{1}{r} \\ \Rightarrow \quad -\frac{G\mu}{r^2} &= -\frac{\ell^2}{\mu^2 r^2}\frac{d^2}{d\phi^2}\frac{1}{r} - \frac{\ell^2}{\mu^2 r^3} \\ \Rightarrow \quad \frac{G\mu^3}{\ell^2} &= \frac{d^2}{d\phi^2}\frac{1}{r} + \frac{1}{r} \end{aligned} \tag{1.57}$$

Solving this differential equation, we get:

$$\begin{aligned} &\frac{1}{r} = (e\cos\phi + 1)\frac{G\mu^3}{\ell^2} \\ \Rightarrow \quad &\boxed{r(\phi) = \left(\frac{\ell^2}{G\mu^3}\right)\left(\frac{1}{1+e\cos\phi}\right)}, \end{aligned} \tag{1.58}$$

for which $e = \sqrt{1 - \ell^2/(aG\mu^3)} \in [0, 1]$, is known as the *eccentricity* and determines the orbit of the masses, with a being the length of the semi-major axis. This thus proves Kepler's first law. The various orbits corresponding to various values of the eccentricity is given as follows:

1. $e = 0$: Circular Orbit
2. $e < 1$: Elliptical Orbit
3. $e = 1$: Parabolic Orbit
4. $e > 1$: Hyperbolic Orbit

Now to prove the third law.

Kepler's Third Law: The square of the planet's orbital period is proportional to the cube of the semi-major axis.

We start by asserting a result we have already found, that is $\frac{dA}{dt} = \text{constant} \equiv h/2$. Since we have an elliptical orbit, we have:

$$\pi ab = \frac{h}{2}T \tag{1.59}$$

where a and b are the semi-major and semi-minor axes respectively, and T is the period of orbit. Let us also define the *perihelion* and *aphelion* of the ellipse as r_1 and r_2. Then we know by geometry of the ellipse that:

$$2a = r_1 + r_2 \tag{1.60}$$

We also use the result from the second law derivation to get:

$$\begin{aligned} & r(0) + r(\pi) = \frac{h^2}{G\mu}\frac{1}{1+e} + \frac{h^2}{G\mu}\frac{1}{1-e} = 2a \\ \Rightarrow\ & a = \frac{h^2}{G\mu}\frac{1}{1-e^2} \\ \Rightarrow\ & b = a\sqrt{1-e^2} \\ \Rightarrow\ & \pi a(a\sqrt{1-e^2}) = \frac{T}{2}\sqrt{G\mu a(1-e^2)} \\ \Rightarrow\ & \frac{T\sqrt{G\mu a}}{2} = \pi a \\ \Rightarrow\ & \boxed{T^2 \propto a^3} \end{aligned} \tag{1.61}$$

Which gives exactly the statement of Kepler's third law.

§1.4.3 Energy Conservation in Central Potential Systems

An interesting direction with which to approach the problem of central potetentials, is via energy conservation. For this, we continue to work in polar coordinates for which the conserved Hamiltonian works out to be

$$\begin{aligned} H &= \dot{r}\frac{\partial \mathcal{L}}{\partial r} + \dot{\phi}\frac{\partial \mathcal{L}}{\partial \dot{\phi}} - \mathcal{L} \\ &= \frac{\mu \dot{r}^2}{2} + \frac{\mu r^2 \dot{\phi}^2}{2} + U(r) \\ &= \frac{\mu \dot{r}^2}{2} + U_{\text{eff}}(r). \end{aligned} \tag{1.62}$$

From this, we get

$$\begin{aligned} & \dot{r} = \pm\sqrt{\frac{2}{\mu}}\sqrt{H - U_{\text{eff}}(r)} \\ \Rightarrow\ & \boxed{\int dt = \pm\sqrt{\frac{\mu}{2}}\int \frac{dr}{\sqrt{H - U_{\text{eff}}(r)}}}. \end{aligned} \tag{1.63}$$

This provides an integral expression that allows us to solve for r as a function of t, and then $\phi(t)$ from the relation $\mu r^2\dot{\phi} = \ell$. Another thing we can do using this integral expression, is to plug into it the angular momentum conservation relation $dt = \mu r^2 d\phi$, which then results in

$$\mu\int d\phi = \pm\sqrt{\frac{\mu}{2}}\int \frac{\ell dr}{r^2\sqrt{H - U_{\text{eff}}(r)}}. \tag{1.64}$$

This will allow us to find ϕ as a function of r (or the inverse), which basically gives us the shape of the trajectory instead of its time trace. For the case of the gravitational potential, solving this equation would result in Eq. (1.58), in which the eccentricity depends on the energy of the system, which determines if the trajectory is closed or not. On the topic of closed orbits, there is a neat theorem regarding this which is presented below.

Theorem 1.4.1. There are only 2 types of central potentials in which all solutions (associated trajectories) are closed. These potentials take the form:

$$U(r) = -\frac{\alpha}{r}, \tag{1.65a}$$

$$U(r) = \alpha r^2, \tag{1.65b}$$

where $\alpha > 0$.

There is a caveat when working with Eq. (1.64), and that it we must require that

$$H > U_{\text{eff}}(r), \tag{1.66}$$

within the bounds of integration, so that the equation is well defined. To ground this intuition a little, we consider the specific potential where $U(r) = -\alpha/r$, which grants that

$$U_{\text{eff}}(r) = \frac{\ell^2}{2\mu r^2} - \frac{\alpha}{r}. \tag{1.67}$$

Potentials of such form look like the illustration in Fig. 1.2.

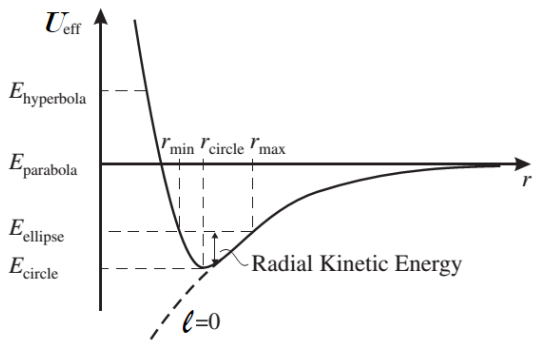

Figure 1.2: Effective potential curve with $U(r) = -\alpha/r$.

We can see that there are several regimes in which the value of E can place the system in, which has been discussed equivalently from the result of Kepler's first law in Sec. 1.4.2 (pertaining to the eccentricity of the orbits). If we consider the energy regime where $E = E_{\text{ellipse}}$, we will retrieve Eq. (1.58) from Eq. (1.64) after evaluating the integrals appropriately. Using the parameters in Fig. 1.2, the shape of the trajectory can also be written as

$$r(\phi) = \frac{A}{1 + e\cos\phi}, \tag{1.68}$$

$$\text{where} \quad A = \frac{x_{\text{max}} + x_{\text{min}}}{2}, \quad e = \frac{x_{\text{max}} - x_{\text{min}}}{x_{\text{max}} + x_{\text{min}}}. \tag{1.69}$$

It is also not too hard to derive for bounded orbits ($E < 0$), the boundary points x_{max} and x_{min}

as

$$x_{\max} = \frac{\mu\alpha}{2} + \sqrt{\frac{\mu^2\alpha^2}{\ell^4} + \frac{2\mu E}{\ell^2}}, \tag{1.70a}$$

$$x_{\min} = \frac{\mu\alpha}{2} - \sqrt{\frac{\mu^2\alpha^2}{\ell^4} + \frac{2\mu E}{\ell^2}}, \tag{1.70b}$$

by setting the denominator in the integrand of Eq. (1.64) to zero. From this, we see that we get circular orbits when:

$$E = -\frac{\mu\alpha^2}{2\ell^2}, \tag{1.71}$$

that results in $e = 0$ and $r = A$. More on the geometry of the ellipse, the center of the ellipse is located a distance

$$d = A\left(\frac{e}{1-e^2}\right), \tag{1.72}$$

from the focus along the line joining the 2 foci. In Cartesian coordinates, the ellipse can be written as

$$\frac{x^2}{a^2} + \frac{(y-d)^2}{b^2} = 1, \tag{1.73}$$

where a is the semi-major axis, b the semi-minor axis and the origin is set at the lower focus with y aligned to the line joining the 2 foci. Explicitly, the semi-major and semi-minor axes can be calculated to be

$$a = \frac{A}{\sqrt{1-e^2}}, \tag{1.74a}$$

$$b = \frac{A}{1-e^2}. \tag{1.74b}$$

For these bounded orbits, one could also ask how long it would take to complete one revolution. This can be easily found from the angular momentum conservation law which gave Eq. (1.59), but written explicitly in terms of our variables defined here, is

$$T = 2\pi\sqrt{\frac{\mu}{\alpha}}\left(\frac{A^{3/2}}{(1-e^2)^{3/2}}\right), \tag{1.75a}$$

$$\text{or} \quad T = \pi\sqrt{\frac{\mu}{2}}\left(\frac{\alpha}{|E|^{3/2}}\right). \tag{1.75b}$$

So we see that the period for an elliptical orbit **only** depends on the energy. As for the unbounded orbits, the expression in Eq. (1.68) shows us that when $E = 0$, we get a parabolic orbit and when $E > 0$, we get hyperbolic orbits.

§1.4.4 The Laplace-Runge-Lenz Vector

Up till now, we have seen that for central potentials systems, they exhibit conservation of energy and angular momentum. However, in the particular case of $U(r) = -\alpha/r$, there is in fact an additional conserved quantity. It also turns out that this conserved quantity is **not** associated to any symmetry of the system and is referred to as the *Laplace-Runge-Lenz vector*

$$\boldsymbol{A} = \boldsymbol{p} \times \boldsymbol{L} - \frac{\mu\alpha}{r}\boldsymbol{r}. \tag{1.76}$$

To check that this is indeed conserved, we can differentiate this quantity with respect to time which gives

$$\begin{aligned} \frac{d}{dt}\boldsymbol{A} &= \frac{d}{dt}\left[\boldsymbol{p} \times \boldsymbol{L} - \frac{\mu\alpha}{r}\boldsymbol{r}\right] \\ &= \frac{d}{dt}\left[\boldsymbol{p}(\boldsymbol{p}\cdot\boldsymbol{r}) - \boldsymbol{r}\boldsymbol{p}^2 - \frac{\mu\alpha}{r}\boldsymbol{r}\right]. \end{aligned} \tag{1.77}$$

From here, we first note that

$$\frac{d\boldsymbol{r}}{dt} = \frac{\boldsymbol{p}}{\mu}, \tag{1.78a}$$

$$\frac{d\boldsymbol{p}}{dt} = -\frac{\alpha}{r^3}\boldsymbol{r}, \tag{1.78b}$$

for which if we plug this into the expression for $\boldsymbol{A}$, gives

$$\frac{d}{dt}\boldsymbol{A} = 0, \tag{1.79}$$

indeed showing that $\boldsymbol{A}$ is conserved. Having the Laplace-Runge-Lenz vector being conserved, it seems as though there would now be a total of $4 + 3 = 7$ conserved quantities which would be physically impossible. It turns out that the Laplace-Runge-Lenz vector only adds **one** additional conserved quantity due to the fact that $\boldsymbol{A}$, is not independent of the other conserved quantities which is made clear by the relations

$$\boldsymbol{A}\cdot\boldsymbol{L} = 0, \tag{1.80a}$$

$$\boldsymbol{A}^2 = \mu^2\alpha^2 + 2\mu L^2 E. \tag{1.80b}$$

These properties also mean that the conservation of $\boldsymbol{A}$ grants us a relationship between ϕ and r, which constrains the trajectories to **periodic, closed orbits**. For clarity, we state this key result below in box.

> 2-Body systems with potentials of the form $U(r) = -\alpha/r$ result in periodic, closed orbits by virtue of the fact that energy E, angular momentum $\boldsymbol{L}$, and the Laplace-Runge-Lenz vector $\boldsymbol{A}$, are conserved.

This result is very counter intuitive because most interacting systems are *ergodic* (all phase space positions will eventually be visited by the trajectories for long enough times). This points to the notion of *integrable systems*, which we will get more into when we study chaos.

§1.4.5 The Spherical Harmonic Oscillator

Another central potential system which results in periodic, closed orbits is the spherically symmetric harmonic oscillator with potential written as

$$U(r) = \alpha r^2, \tag{1.81}$$

where $\alpha > 0$. In this system, we can write out the Lagrangian explicitly in Cartensian coordinates as

$$\begin{aligned} \mathcal{L} &= \frac{\mu}{2}\left(\dot{x}^2 + \dot{y}^2 + \dot{z}^2\right) - \alpha(x^2 + y^2 + z^2) \\ &= \left(\frac{\mu\dot{x}^2}{2} - \alpha x^2\right) + \left(\frac{\mu\dot{y}^2}{2} - \alpha y^2\right) + \left(\frac{\mu\dot{z}^2}{2} - \alpha z^2\right). \end{aligned} \tag{1.82}$$

Writing the Lagrangian in this form clearly shows that it factorizes into its 3 non-interacting axes, which would result in independent energy conservation for each axis $E_x, E_y, E_z =$ constant, along with conservation of angular momentum $\boldsymbol{L}$. However, these conserved quantities are **not** all independent of each other.

Chapter 2

Classical Scattering

Scattering is the study of interacting particles which approach each other from infinity, interact due to a finite-range potential and then fly off into infinity once more. The focus of such a study is determining the resultant directions of scatterers after a scattering event has occured, given the incoming scattering directions are known. Scattering theory is a powerful tool for us to infer properties of the scatterers via performing such scattering experiments, and is very useful in quantum mechanical and particle physics systems.

§2.1 Introduction

Consider a heavy object, which we will call the *scattering center*, located at some arbitrary point in space. Now imagine that an experimentalist throws (in the same manner) several smaller objects, which we will call *scatterers*, at this scattering center with masses such that the scattering center remains effectively stationary. After interacting with the scattering center and flying off, we can then count the fraction of scatterers which fly off into a particular *solid angle*, Ω (the angular area on an imaginary spherical surface surrounding the scattering center).

From this fraction, we can define a quantity that we will use to characterize the scattering center, known as the *scattering cross section*, Σ which is defined as the ratio of the area through which the scatterers are thrown at the scattering center A, to the solid angle, Ω

$$\Sigma = \frac{A}{\Omega}. \tag{2.1}$$

Differential analysis can be applied as well to such problems in the case where a more fine-grained description of the scattering center is required, for which we define an analogous quantity known as the *differential cross section*

$$\Sigma_{\text{diff}} = \frac{dA}{d\Omega}, \tag{2.2}$$

often denoted as $\Sigma_{\text{diff}} = d\sigma/d\Omega$. From this, the *total scattering cross section* is then be computed as

$$\boxed{\sigma_{\text{total}} = \int \frac{d\sigma}{d\Omega} d\Omega}. \tag{2.3}$$

A visualization of the set-up for such a problem is provided in Figs. 2.1 and 2.2.

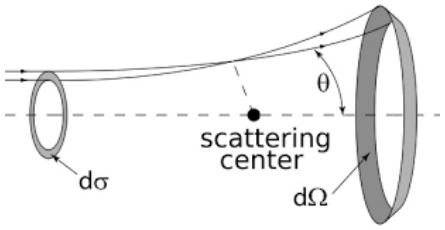

Figure 2.1: Set-up of scattering problems.

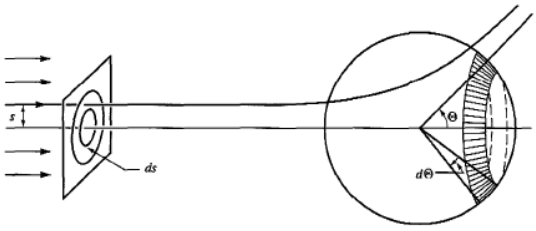

Figure 2.2: Set-up of scattering problems (Goldstein).

Formally, we can describe the interaction between the scatterers and scattering center with a potential function $U(\boldsymbol{r})$. It would be appropriate now to introduce some additional parameters that will help to characterize the problem:

1. Impact parameter, s – the orthogonal distance from the principle axis to the incoming differential scattering area, $d\sigma$;
2. Scattering angles, (θ, ϕ) – the angle between the principle axis and the differential solid angle $d\Omega$ into which scatterers are scattered, with origin at the scattering center.

A simplification that can be made to the problem is to once again assert that $U(\boldsymbol{r}) = U(r)$, which causes ϕ to drop out of the problem. Hence, only θ is required to specify the scattered system, which allows us to compute the differential cross section as

$$dA = 2\pi s ds \tag{2.4}$$

$$d\Omega = 2\pi \sin\theta d\theta \tag{2.5}$$

$$\Rightarrow \quad \boxed{\frac{d\sigma}{d\Omega} = \frac{s ds}{\sin\theta d\theta} = \frac{1}{2\sin\theta}\frac{d(s^2)}{d\theta}}. \tag{2.6}$$

The impact parameter s, can be thought of a variable that parameterizes the angular momentum of the incoming scatterers, $\ell = s \times (\mu v_{\text{in}})$, where v_{in} is the incoming velocity.

§2.1.1 Rutherford Scattering

A famous case of scattering is that of *Rutherford scattering*. Rutherford scattering is the case of scattering due to an electric Coulomb potential

$$U(r) = -\frac{Q_1 Q_2}{4\pi\varepsilon_0 r}. \tag{2.7}$$

For simplicity, we will again group all factors in the potential into the constant α with units of energy $\times$ length. Such an interaction potential has already been studied in the previous chapter, so we will be taking a lot of the results we got from there for this problem. More specifically, we are concerned with the case where the incoming scatterer is **not** "captured" by the attractive potential of the scattering center, but rather follows an unbounded trajectory. We know that for this potential, we have that the trajectories follow

$$r(\phi) = \frac{A}{1 + e\cos\phi}, \tag{2.8}$$

where $e > 1$ for unbounded (hyperbolic) trajectories. Explicitly, the eccentricity is given by the formula

$$e = \sqrt{1 + \frac{2E\ell^2}{\mu\alpha^2}} = \sqrt{1 + \frac{4E^2 s^2}{\alpha^2}}, \tag{2.9}$$

where $E = (\ell/s)^2/(2\mu)$ is the energy of the incoming particle. From the geometry of the problem, we will have the scattering angle corresponding to 2 times the angle of asymptotic trajectory

$$\phi_{\text{asymp}} = \arcsin\left(\frac{1}{e}\right) \tag{2.10}$$

$$\Rightarrow \quad \theta = 2\arcsin\left(\frac{1}{e}\right). \tag{2.11}$$

As such, we can substitute the expression for the eccentricity into the equation above to get

$$\begin{aligned} &\sin\left(\frac{\theta}{2}\right) = \frac{1}{\sqrt{1 + \frac{4E^2 s^2}{\alpha^2}}} \\ \Rightarrow \quad &s^2 = \frac{\alpha^2}{4E^2}\cot^2\left(\frac{\theta}{2}\right) \\ \Rightarrow \quad &\boxed{\frac{d\sigma}{d\Omega} = \frac{1}{2\sin\theta}\frac{d(s^2)}{d\theta} = \frac{\alpha^2}{16E^2}\frac{1}{\sin^4\left(\frac{\theta}{2}\right)}}. \end{aligned} \tag{2.12}$$

Interestingly, this result when derived from quantum mechanics works out to be exactly the same despite this being a classical derivation!

Note: Notice that when the scattering angle goes to zero, the differential cross section diverges, implying strong forward scattering. Furthermore, the total cross section also diverges, which implies that these Coulomb potentials have an extremely strong tendency to scatter.

§2.2 Galilean Transformations

Galilean transformations are coordinate transformations in which classical mechanics remains invariant under their action (i.e. the action remains unaffected by these transformations). The most common of these transformations is known as a *Galilean boost* $\mathcal{B}$, defined as

$$\boldsymbol{x}' = \mathcal{B}[\boldsymbol{x}] = \boldsymbol{x} - \boldsymbol{v}t, \tag{2.13}$$

where $\boldsymbol{v}$ is the velocity in which the unprimed frame moves relative to the primed frame. To see that classical mechanics indeed remains invariant under this transformation, we apply this

to the action ($S[\boldsymbol{x}] \to S[\mathcal{B}[\boldsymbol{x}]]$), which gives

$$\begin{aligned} S[\mathcal{B}[\boldsymbol{x}]] &= \int dt\mathcal{B}\left[\frac{1}{2}\sum_j m_j\dot{\boldsymbol{x}}_j^2 - \sum_{i<j} U(\boldsymbol{x}_i - \boldsymbol{x}_j)\right] \\ &= \int dt\left[\frac{1}{2}\sum_j m_j(\dot{\boldsymbol{x}}_j - \boldsymbol{v})^2 - \sum_{i<j} U(\boldsymbol{x}_i - \boldsymbol{x}_j)\right] \\ &= S[\boldsymbol{x}_j] + \int dt\sum_j m_j\left[\frac{1}{2}\boldsymbol{v}^2 - \dot{\boldsymbol{x}}_j\boldsymbol{v}\right] \\ &= S[\boldsymbol{x}_j] + \int dt\sum_j \frac{1}{2}m_j\boldsymbol{v}^2 - \sum_j (\boldsymbol{x}_j\boldsymbol{v})\Big|_{t_i}^{t_f}. \end{aligned} \tag{2.14}$$

Notice that the second term in the last line above, has simply an integrand which is a constant and will thus not affect the equations of motion. Furthermore, the last term in the last line above is only dependent on the end points of the trajectories, which do not get varied in the action. This leaves us with

$$S[\mathcal{B}[\boldsymbol{x}_j]] = S[\boldsymbol{x}_j], \tag{2.15}$$

indeed showing that this transformation leaves the physics unchanged. Let us now see an implication of this in the context of scattering. Consider again the scattering of 2 classical particles, for which we adopt center of mass coordinates

$$\boldsymbol{r} = \boldsymbol{r}_2 - \boldsymbol{r}_1, \tag{2.16a}$$

$$\boldsymbol{R} = \frac{m_1\boldsymbol{r}_1 + m_2\boldsymbol{r}_2}{m_1 + m_2}. \tag{2.16b}$$

Because there are not net forces on the center of mass of the 2-body system, we have that $\dot{\boldsymbol{R}} \equiv \boldsymbol{V}$, remains constant throughout the motion. As such, we can consider a Galilean boost into a frame which moves with the center of mass of the system. In this boosted frame, we will have the velocities as

$$\boldsymbol{v}_1' = \boldsymbol{v}_1 - \boldsymbol{V} = \frac{m_2}{m_1 + m_2}(\boldsymbol{v}_1 - \boldsymbol{v}_2), \tag{2.17a}$$

$$\boldsymbol{v}_2' = \boldsymbol{v}_2 - \boldsymbol{V} = -\frac{m_1}{m_1 + m_2}(\boldsymbol{v}_1 - \boldsymbol{v}_2). \tag{2.17b}$$

Since the center of mass and relative motion completely decouples, we have that the magnitude of the relative velocity will remain constant ($v = |\boldsymbol{v}_1 - \boldsymbol{v}_2| = \text{constant}$). This indicates that after scattering (denoted with tildes), the velocities would be

$$\tilde{\boldsymbol{v}}_1' = \frac{m_2}{m_1 + m_2}v\hat{\boldsymbol{n}}, \tag{2.18a}$$

$$\tilde{\boldsymbol{v}}_2' = -\frac{m_1}{m_1 + m_2}v\hat{\boldsymbol{n}}, \tag{2.18b}$$

where $\hat{\boldsymbol{n}}$ is a unit vector denoting the post-collision scattering direction. As such, we will have that the final velocities are written as

$$\boxed{\tilde{\boldsymbol{v}}_1 = \boldsymbol{V} + \frac{m_2}{m_1 + m_2}v\hat{\boldsymbol{n}}, \quad \tilde{\boldsymbol{v}}_2 = \boldsymbol{V} - \frac{m_1}{m_1 + m_2}v\hat{\boldsymbol{n}}}. \tag{2.19}$$

§2.3 Virial Theorem

There is a unique feature that arises for systems under a potential which polynomially scales with its coordinates. That is to say, for a given scaling of a coordinate q, by a constant λ, the potential is modified by

$$q \quad \rightarrow \quad q' = \lambda q \tag{2.20}$$

$$\Rightarrow \quad U(q) \quad \rightarrow \quad U(q') = \lambda^{\beta} U(q), \tag{2.21}$$

where β is the integer associated to the power of the coordinate in the potential (e.g. if $U(r) = -\alpha/r$, then $\beta = -1$). We know that the kinetic energy of the system scales with velocity by a square relation

$$T(v)' = T(\lambda v) = \lambda^2 T(v), \tag{2.22}$$

which allows us to write the relation

$$\sum_j \boldsymbol{v}_j \cdot \frac{\partial T}{\partial v_j} = 2T. \tag{2.23}$$

Consider a system that undergoes closed periodic orbits, i.e., systems in which $\boldsymbol{r}(t) = \boldsymbol{r}(t+T)$ and $\dot{\boldsymbol{r}}(t) = \dot{\boldsymbol{r}}(t+T)$. In such integrable systems, we can consider time-averages of quantities such as energy, over the periodic motion which we shall denote

$$\langle \ldots \rangle = \frac{1}{T} \int_0^T (\ldots) dt. \tag{2.24}$$

At this point, we consider the following series of equalities regarding kinetic energy T, and the associated generalized coordinates (along with their time derivatives):

$$\begin{aligned} 2T &= \sum_j \boldsymbol{v}_j \cdot \frac{\partial T}{\partial v_j} \\ &= \sum_j \boldsymbol{p}_j \cdot \boldsymbol{v}_j \\ &= \sum_j \left[\frac{d}{dt} \left(\boldsymbol{p}_j \cdot \boldsymbol{r}_j \right) - \dot{\boldsymbol{p}}_j \cdot \boldsymbol{r}_j \right] \\ &= \sum_j \left[\frac{d}{dt} \left(\boldsymbol{p}_j \cdot \boldsymbol{r}_j \right) + \frac{\partial V}{\partial \boldsymbol{r}_j} \cdot \boldsymbol{r}_j \right]. \end{aligned} \tag{2.25}$$

Now time-averaging these quantities gives

$$2\langle T \rangle = \sum_j \left\langle \boldsymbol{\nabla}_{\boldsymbol{r}_j} V \cdot \boldsymbol{r}_j \right\rangle, \tag{2.26}$$

where we note that the total time derivative term drops out in the time-averaging over one period, by virtue of the trajectory being periodic. If we now assert that the potential has polynomial scaling with its coordinates, we then have the relation

$$\boxed{2\langle T \rangle = \beta \langle U \rangle}. \tag{2.27}$$

This result is known as the *Virial theorem*.

Example:

Consider a planetary system in which a planet undergoes closed periodic motion around a star (gravitational potential energy would leave us with $\beta = -1$). If we want to know what the total energy of the system is, we can simply employ the Virial theorem, knowing that

$$\begin{aligned} & U = -\frac{GMm}{R} \\ \Rightarrow \quad & E = T + U \\ & = -\frac{1}{2}U + U \\ & = -\frac{GMm}{2R}. \end{aligned} \tag{2.28}$$

Chapter 3

Rigid Body Motion

The study of rigid bodies dynamics presents a vast landscape of rich phenomena, which on its own could constitute stand alone books just by the sheer expanse of its complexity. In this chapter, we aim to study the basics and several key concepts of this topic, with no attempt to be completely comprehensive.

§3.1 Introduction

We start with the definition of what a rigid body is.

Definition 3.1.1. Rigid Bodies: *A rigid body is defined by a collection of points, in which their relative coordinate configuration is preserved throughout its dynamical evolution.*

With just this notion of a rigid body, the dynamics of such extended objects can in principle get arbitrarily complicated. Thankfully, the work of Euler allowed a simplification to such systems (at least in terms of approach), for which his theorem states that arbitrary motion of a rigid body can always be decomposed into

1. translational motion of the rigid body's center of inertia;
2. rotational motion of the rigid body about its center of inertia.

Note: The point about which translation and rotation is defined does **not** necessarily have to be the center of inertia, but it is often convenient to specify it as such.

Let's first further delve into the subject of rotations.

§3.1.1 Rigid Body Rotations

Rotations of a rigid body can be characterized by specifying an axis $\hat{\boldsymbol{n}}$, through which the rotation occurs, and an angle θ, by which the objects rotates. This formalism is known as the *axis-angle* formulation of rotations and conventionally follows the right-handed rotation convention. The 2

pieces of information can be grouped into a single vector we call $\boldsymbol{\varphi}$, which points in the direction $\hat{\boldsymbol{n}}$ and has magnitude θ:

$$\boldsymbol{\varphi} = \theta\hat{\boldsymbol{n}}. \tag{3.1}$$

With this, we can that for an infinitesimal rotation, each point on the rigid body specified by coordinate $\boldsymbol{r}_k$, relative to the center of inertia moves by an amount

$$d\boldsymbol{r}_k = d\boldsymbol{\varphi} \times \boldsymbol{r}_k = \hat{\boldsymbol{n}}d\theta \times \boldsymbol{r}_k. \tag{3.2}$$

Adding in the translational motion of the center of mass $d\boldsymbol{R}$, we get that the total motion of a rigid body can be written as

$$\boxed{d\boldsymbol{r}_k = d\boldsymbol{R} + (d\boldsymbol{\varphi} \times \boldsymbol{r}_k)}. \tag{3.3}$$

From this expression, we can derive the velocity of the points as

$$\frac{d\boldsymbol{r}_k}{dt} = \frac{d\boldsymbol{R}}{dt} + \left(\frac{d\boldsymbol{\varphi}}{dt} \times \boldsymbol{r}_k\right), \tag{3.4}$$

which can be more succinctly written as

$$\boxed{\boldsymbol{v}_k = \boldsymbol{V} + (\boldsymbol{\Omega} \times \boldsymbol{r}_k)}. \tag{3.5}$$

where $\boldsymbol{v}_k = d\boldsymbol{r}_k/dt$, $\boldsymbol{V} = d\boldsymbol{R}/dt$ and $\boldsymbol{\Omega} = d\boldsymbol{\varphi}/dt$. The main application of this formula, is to compute the kinetic energy of the system which can then be placed into the Lagrangian to derive the equations of motion.

For the sake of curiosity, let's say that we had not chosen the center of inertia to base translation and rotation of the object around. We instead chose another point such that to move from the center of inertia frame to this newly defined frame (denoted by tildes), we have to translate by a vector $\boldsymbol{a}$ such that

$$\tilde{\boldsymbol{r}} = \boldsymbol{r} + \boldsymbol{a}. \tag{3.6}$$

Plugging these back into the expression we have above, we get

$$\begin{aligned} d\tilde{\boldsymbol{r}}_k &= d\boldsymbol{R} + [d\boldsymbol{\varphi} \times (\tilde{\boldsymbol{r}}_k - \boldsymbol{a})] \\ &= [d\boldsymbol{R} - (d\boldsymbol{\varphi} \times \boldsymbol{a})] + (d\boldsymbol{\varphi} \times \tilde{\boldsymbol{r}}_k) \\ &= d\tilde{\boldsymbol{R}} + (d\boldsymbol{\varphi} \times \tilde{\boldsymbol{r}}_k). \end{aligned} \tag{3.7}$$

Dividing through by dt then brings us back to

$$\frac{d\tilde{\boldsymbol{r}}_k}{dt} = \frac{d\tilde{\boldsymbol{R}}}{dt} + (\boldsymbol{\Omega} \times \tilde{\boldsymbol{r}}_k), \tag{3.8}$$

so we see that the motion can still be described in the same way, and so the definition of the point of reference does not affect the physics, just the mathematical convenience of working in those coordinates.

§3.2 Rigid Body Dynamics

Now having the form of the velocity of each point in the rigid body, we can write down its total kinetic energy as

$$\begin{aligned} T &= \sum_k \frac{m_k}{2} \left[\boldsymbol{V} + (\boldsymbol{\Omega} \times \boldsymbol{r}_k)\right]^2 \\ &= \sum_k \frac{m_k}{2} \boldsymbol{V}^2 + \sum_k m_k \boldsymbol{V} \cdot (\boldsymbol{\Omega} \times \boldsymbol{r}_k) + \sum_k \frac{m_k}{2} \left(\boldsymbol{\Omega} \times \boldsymbol{r}_k\right)^2 \end{aligned} \tag{3.9}$$

At this point, we can consider the 3 terms above and simplify them individually by first introducing some definitions.

> **Note**: For generality, we will define the coordinate frame we attach to the rigid body as having an origin at an arbitrary point $\boldsymbol{R}$, on it (not necessarily at the center of inertia).

We define $M = \sum_k m_k$ as the total mass and $\boldsymbol{\rho} = \sum_k m_k \boldsymbol{r}_k$ as the vector connecting $\boldsymbol{R}$ to the rigid body's center of mass, which will allow us to have the 3 terms written as:

1. First term:

$$\sum_k \frac{m_k}{2} \boldsymbol{V}^2 = \frac{M}{2} \boldsymbol{V}^2, \tag{3.10}$$

   which is a center of mass kinetic energy;

2. Second term:

$$\begin{aligned} \sum_k m_k \boldsymbol{V} \cdot (\boldsymbol{\Omega} \times \boldsymbol{r}_k) &= \boldsymbol{V} \cdot \left(\boldsymbol{\Omega} \times \sum_k m_k \boldsymbol{r}_k\right) \\ &= M \boldsymbol{V} \cdot (\boldsymbol{\Omega} \times \boldsymbol{\rho}) \\ &= \boldsymbol{\Omega} \cdot (\boldsymbol{\rho} \times M \boldsymbol{V}) \\ &= \boldsymbol{\Omega} \cdot \boldsymbol{L}_{\mathrm{CoM}}, \end{aligned} \tag{3.11}$$

   which is a projection of the angular momenta of the rigid body due to its center of mass about $\boldsymbol{R}$ and that about its center of mass itself. We see that term immediately vanishes if we are in the center of mass frame ($\boldsymbol{\rho} = 0$), or there is a rigid attachment such that $\boldsymbol{V} = 0$;

3. Third term:

$$\begin{aligned}\sum_k \frac{m_k}{2}\left(\boldsymbol{\Omega}\times\boldsymbol{r}_k\right)^2 &= \sum_k \frac{m_k}{2}\left[\boldsymbol{\Omega}^2\boldsymbol{r}_k^2 - \left(\boldsymbol{\Omega}\cdot\boldsymbol{r}_k\right)^2\right] \\ &= \sum_{m,n}\frac{1}{2}\Omega_m\Omega_n\sum_k m_k\delta_{m,n}\left(\boldsymbol{r}_k^2 - r_k^m r_k^n\right) \\ &= \frac{1}{2}\sum_{m,n} I_{m,n}\Omega_m\Omega_n \\ &= \frac{1}{2}\boldsymbol{\Omega}^T\boldsymbol{I}\boldsymbol{\Omega},\end{aligned} \tag{3.12}$$

which is the rotational kinetic energy about the chosen point in the body, where

$$\boxed{I_{i,j} = \sum_k m_k\left(\boldsymbol{r}_k^2 - r_k^i r_k^j\right)\delta_{i,j}} \tag{3.13}$$

$$= \int d^3r\rho(\boldsymbol{r})\left(\boldsymbol{r}^2 - r_i r_j\right)\delta_{i,j}, \qquad \text{(for continuous bodies)} \tag{3.14}$$

is known as the *moment of inertia tensor*. This tensor is symmetric under exchange of indices and can be used to write down the angular momentum of the system as

$$\boldsymbol{L} = \boldsymbol{I}\boldsymbol{\Omega}. \tag{3.15}$$

Note: The body-frame coordinate axes in which the moment of inertia tensor becomes diagonal are known as the *principal axes of inertia*, which have diagonal elements referred to as the *principal moments of inertia*. In the diagonal basis, we will always have the inequality relation $I_{11} \leq I_{22} + I_{33}$ by construction of this tensor (equality occurs only for planar bodies rotating in the plane of rotation).

In summary, we have that the kinetic energy term can be written as

$$\boxed{T = \frac{1}{2}M\boldsymbol{V}^2 + \boldsymbol{\Omega}\cdot\boldsymbol{L}_{\text{CoM}} + \frac{1}{2}\boldsymbol{\Omega}^T\boldsymbol{I}\boldsymbol{\Omega}}. \tag{3.16}$$

In general, the difficulty in solving for the motion of rigid bodies primarily arises from the moment of inertia tensor, which in practice could be difficult to compute for arbitrary rigid bodies. In situations where rigid body possesses a degree of symmetry (e.g. spherical symmetry, cylindrical symmetry, etc...), this could simplifies the problem as the axis of symmetry would be indicative of where the principle axis of inertia would lie. Otherwise, an informed guess or explicitly term-by-term calculation of the moment if inertia tensor would be necessary. Thankfully, a theorem does exist which aids in the approach to problems of more complex geometries. This is known as the *parallel axis (Huygens–Steiner) theorem*:

Theorem 3.2.1. *Given that the moment of inertia tensor about a rigid body's center of mass I_{ij}, through some axis $\hat{\boldsymbol{n}}$ is known, the moment of inertia about a different point $\tilde{I}_{ij}$, separated by vector $\boldsymbol{a}$ through an axis parallel to $\hat{\boldsymbol{n}}$ is given by:*

$$\tilde{I}_{ij} = I_{ij} + M\delta_{i,j}\left(\boldsymbol{a}^2 - a_i a_j\right). \tag{3.17}$$

Note that this only works when the moment of inertia about the ***center of mass*** *is known.*

We will come back to using theorem. 3.2.1 later on in the chapter.

§3.2.1 Equations of Motion for Free Rigid Bodies

Having the form of the kinetic energy in Eq. (3.16), we can now consider the equations of motion for a free rigid body (under no influence of external forces) via the Euler-Lagrange equations. To start, we assert that the reference point passes through the rigid body's center of mass, and so the Lagrangian is simply given by

$$\mathcal{L} = T = \frac{1}{2}M\boldsymbol{V}^2 + \frac{1}{2}\boldsymbol{\Omega}^T\boldsymbol{I}\boldsymbol{\Omega}. \tag{3.18}$$

A free body will have angular momentum conserved, so we have

$$L_j = \frac{\partial \mathcal{L}}{\partial \Omega_j} = \sum_k I_{j,k}\Omega_k = \text{constant}. \tag{3.19}$$

For simplicity, we start by considering a cylindrically symmetric object, such that its principal moments obey $I_1 = I_2 \neq I_3$. Such objects are known as *symmetric tops*, where I_{33} is the principal moment along the axis of cylindrical symmetry and I_{11} and I_{22} corresponding to axes orthogonal to this. We now consider the case in which the symmetric top rotates about axis 3, and also has motion about an axis $\hat{\boldsymbol{n}}$, which makes an angle θ from the axis of cylindrical symmetry (axis 3) with angular momentum $\boldsymbol{L}$. We can define our lab-frame such that $\hat{\boldsymbol{n}} = \hat{z}$, and consider the projection of the angular momentum onto the principal axes, which gives

$$L\cos\theta = I_3\Omega_3 \tag{3.20}$$

$$L\sin\theta = I_1\Omega_1 = I_2\Omega_2. \tag{3.21}$$

This system is visualized in Fig. 3.1.

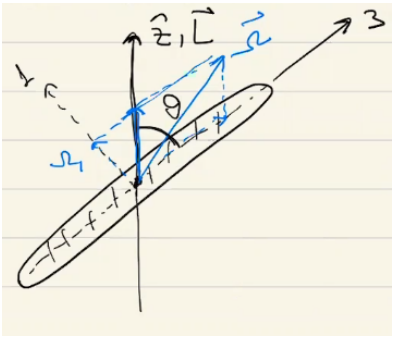

Figure 3.1: Rotating symmetric top.

In the lab-frame the moment of inertia tensor is not diagonal, so the angular velocity vector $\boldsymbol{\Omega}$ would not (in general) point along either $\hat{z}$ or axis 3. It would however, lie in the plane defined by $\hat{z}$ and axis 3. From here, a quantity we would like to know is known as the *precession frequency* Ω_p, which is how fast axis 3 rotates about $\hat{z}$. To compute this, we projection the angular momentum onto axis 1, which grants the relation

$$\begin{aligned} \frac{L\sin\theta}{I_1} &= \Omega_p \sin\theta \\ \Rightarrow \quad \Omega_p &= \frac{L}{I_1}. \end{aligned} \tag{3.22}$$

In a general asymmetric free rigid body ($I_1 \neq I_2 \neq I_3$), the precession frequency about the z-axis can be found by projecting the motion of ϕ on to the x_1, x_2-plane, which gives the relation

$$\begin{aligned} \dot{\phi}\sin\theta &= \Omega_1\cos\psi + \Omega_2\sin\psi \\ \Rightarrow \quad \dot{\phi} &= \frac{\Omega_1\cos\psi + \Omega_2\sin\psi}{\sin\theta}. \end{aligned} \tag{3.23}$$

Defining the conserved angular momentum along z as $\boldsymbol{\ell}$, we would then have the relations

$$I_1\Omega_1 = \ell\sin\theta\cos\psi, \tag{3.24a}$$
$$I_2\Omega_2 = \ell\sin\theta\sin\psi, \tag{3.24b}$$

which can be plugged into the expression for $\dot{\phi}$ to give

$$\dot{\phi} = \frac{I_1\Omega_1^2 + I_2\Omega_2^2}{\ell\sin^2\theta} = \ell\left(\frac{I_1\Omega_1^2 + I_2\Omega_2^2}{I_1^2\Omega_1^2 + I_2^2\Omega_2^2}\right). \tag{3.25}$$

§3.2.2 Torques on Rigid Bodies

We now consider a rigid body under the influence of some potential U, such that the Lagrangian is given by

$$\begin{aligned} \mathcal{L} &= T - U \\ &= \frac{1}{2}M\boldsymbol{V}^2 + \frac{1}{2}\boldsymbol{\Omega}^T\boldsymbol{I}\boldsymbol{\Omega} - \sum_k U_k. \end{aligned} \tag{3.26}$$

Since the potential term will be dependent only on the real-space coordinates, we can consider the change in the potential under an infinitesimal rotation of the rigid body. This is given by

$$\begin{aligned} U_{k'} &= U(\boldsymbol{r}_k') \\ &= U(\boldsymbol{R} + d\boldsymbol{\varphi}\times\boldsymbol{r}_k) \\ &\approx U(\boldsymbol{R}) + \boldsymbol{\nabla}_{\boldsymbol{r}_k}U\cdot(d\boldsymbol{\varphi}\times\boldsymbol{r}_k) \\ &= U(\boldsymbol{R}) - \boldsymbol{F}_k\cdot(d\boldsymbol{\varphi}\times\boldsymbol{r}_k) \\ &= U(\boldsymbol{R}) - d\boldsymbol{\varphi}\cdot(\boldsymbol{r}_k\times\boldsymbol{F}_k) \\ &= U(\boldsymbol{R}) - d\boldsymbol{\varphi}\cdot\boldsymbol{\tau}_k, \end{aligned} \tag{3.27}$$

where $\boldsymbol{\tau} \equiv \boldsymbol{r} \times \boldsymbol{F}$ is the *torque*. If we plug this into the Euler-Lagrange equations, we end up with the rotational analog to Newton's second law:

$$\boxed{\frac{d}{dt}\boldsymbol{L} = \sum_k \boldsymbol{\tau}_k = \sum_k (\boldsymbol{r}_k \times \boldsymbol{F}_k)}. \tag{3.28}$$

This is not to be confused with the resultant force on a rigid body which can be written as the sum over forces on all its constituent parts

$$M\dot{\boldsymbol{V}} = -\sum_k \frac{\partial U_k}{\partial \boldsymbol{r}_k} = \sum_k \boldsymbol{F}_k. \tag{3.29}$$

§3.2.3 Euler Angles

In 3 dimensional space, we will require 3 angular degrees of freedom to fully specify the change in orientation of a rigid body. A formalism for this was introduced by Euler known as the *Euler angles*. The angles allow for a means of decomposition for the rotation undergone by a rigid body. These angles specify the following rotational actions on an initial Cartesian frame:

1. θ: a rotation about the x-axis, $(x, y, z) \to (x, y', z')$;
2. ϕ: a rotation about the z-axis, $(x, y', z') \to (x', y'', z')$;
3. ψ: a rotation about the z'-axis, $(x', y'', z') \to (x'', y'', z')$,

where all rotation angles follow the right-handed convention. This is visualized in Fig. 3.2.

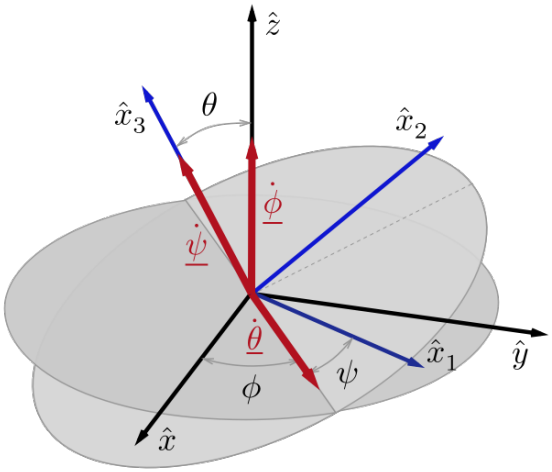

Figure 3.2: Visualization of Euler angles.

It is then possible to derive relations for the angular speed of the principal axes in terms of the Euler angles, which are written as

$$\Omega_1 = \dot{\phi}\sin\theta\sin\psi + \dot{\theta}\cos\psi, \tag{3.30a}$$
$$\Omega_2 = \dot{\phi}\sin\theta\cos\psi - \dot{\theta}\sin\psi, \tag{3.30b}$$
$$\Omega_3 = \dot{\phi}\cos\theta + \dot{\psi}. \tag{3.30c}$$

Note: The relations in Eqs. (3.30a) describe the motion of a rigid body described by coordinate frame (x_1, x_2, x_3) relative to the coordinate frame (x, y, z).

Physically, $\dot{\phi}$ corresponds to the precession of axis 3 about the z-axis. The resulting kinetic energy of the body in the body-frame can then be written in terms of these Euler angles as

$$\begin{aligned} T &= \frac{1}{2}\sum_{j=1}^{3} I_j \Omega_j^2 \\ &= \frac{1}{2}\left[I_1 \left(\dot{\phi}\sin\theta\sin\psi + \dot{\theta}\cos\psi \right)^2 + I_2 \left(\dot{\phi}\sin\theta\cos\psi - \dot{\theta}\sin\psi \right)^2 + I_3 \left(\dot{\phi}\cos\theta + \dot{\psi} \right)^2 \right]. \end{aligned} \tag{3.31}$$

This kinetic energy greatly simplifies for symmetric tops ($I_1 = I_2 \neq I_3$), to give

$$T = \frac{1}{2} I_1 \left(\dot{\phi}^2 \sin^2\theta + \dot{\theta}^2 \right) + \frac{1}{2} I_3 \left(\dot{\psi} + \dot{\phi}\cos\theta \right)^2. \tag{3.32}$$

§3.2.4 Euler's Equations

The Lagrangian for asymmetric rigid bodies in terms of Euler angles are pretty daunting (as seen above) to simply plug into the Euler-Lagrange equations. However, Euler also derived another set of equations known as the *Euler equations* which simplify the problem for a rotating rigid-body. The way to do this is to move into the *rotating frame* of the rigid body, which results in any vector quantity $\boldsymbol{A}$ to vary as

$$\frac{d\boldsymbol{A}}{dt} = \left.\frac{d\boldsymbol{A}}{dt}\right|_{\text{R.F.}} + (\boldsymbol{\Omega} \times \boldsymbol{A}), \tag{3.33}$$

where $d\boldsymbol{A}/dt|_{\text{R.F.}}$ is the change of the vector $\boldsymbol{A}$ inside the rotating frame. In the case where $\boldsymbol{A} = \boldsymbol{L}$, the angular momentum of the rigid body, we have

$$\frac{d\boldsymbol{L}}{dt} = \left.\frac{d\boldsymbol{L}}{dt}\right|_{\text{R.F.}} + (\boldsymbol{\Omega} \times \boldsymbol{L}) = \begin{bmatrix} I_1\dot{\Omega}_1 + \Omega_2\Omega_3(I_3 - I_2) \\ I_3\dot{\Omega}_2 + \Omega_3\Omega_1(I_1 - I_3) \\ I_3\dot{\Omega}_3 + \Omega_1\Omega_2(I_2 - I_1) \end{bmatrix}. \tag{3.34}$$

If there are net torques that act on the body, we then have to equation this to $d\boldsymbol{L}/dt$, which results in the Euler equations:

$$\boxed{\begin{aligned} I_1\dot{\Omega}_1 + \Omega_2\Omega_3(I_3 - I_2) &= \sum_k \tau_1^k, \\ I_2\dot{\Omega}_2 + \Omega_3\Omega_1(I_1 - I_3) &= \sum_k \tau_2^k, \\ I_3\dot{\Omega}_3 + \Omega_1\Omega_2(I_2 - I_1) &= \sum_k \tau_3^k. \end{aligned}} \tag{3.35}$$

Example:

Consider a free symmetric top, such that $I_1 = I_2 \neq I_3$ and $\boldsymbol{\tau}_k = 0$ for all k. As such, the

Euler equations grant us that

$$I_1\dot{\Omega}_1 + \Omega_2\Omega_3(I_3 - I_2) = 0, \tag{3.36a}$$
$$I_2\dot{\Omega}_2 + \Omega_3\Omega_1(I_1 - I_3) = 0, \tag{3.36b}$$
$$I_3\dot{\Omega}_3 = 0. \tag{3.36c}$$

As such, we get that

$$\Omega_3 = \text{constant}, \tag{3.37a}$$
$$\dot{\Omega}_1 = -\omega\Omega_2, \tag{3.37b}$$
$$\dot{\Omega}_2 = +\omega\Omega_1, \tag{3.37c}$$

where

$$\omega \equiv \left(\frac{I_3 - I_1}{I_1}\right)\Omega_3. \tag{3.38}$$

The time dependence of Ω_1 and Ω_2 is obtained by solving the coupled differential equations to get

$$\Omega_1 = \Omega_0\cos(\omega t), \quad \Omega_2 = \Omega_0\sin(\omega t), \tag{3.39}$$

where Ω_0 is some constant dependent on initial conditions. This result is exactly equivalent to that obtained in Sec. 3.2.1.

§3.2.5 Euler Equations for Free Rigid Bodies

In the absence of external forces and torques, a rigid body would have conservation of energy and angular momentum

$$E = \frac{1}{2}\left(I_1\Omega_1^2 + I_2\Omega_2^2 + I_3\Omega_3^2\right) = \text{constant}, \tag{3.40a}$$
$$\ell^2 = I_1^2\Omega_1^2 + I_2^2\Omega_2^2 + I_3^2\Omega_3^2 = \text{constant}, \tag{3.40b}$$

which largely constrains the complexity of the problem. The result of this is that there is only 1 independent angular momentum Ω_j (i.e. the other 2 angular momenta can be written in terms of Ω_j). Picking $j = 3$, we have

$$\Omega_1^2(\Omega_3) = \frac{I_3(I_2 - I_3)\Omega_3^2 + \ell^2 - 2I_2E}{I_1(I_1 - I_2)}, \tag{3.41a}$$
$$\Omega_2^2(\Omega_3) = \frac{I_3(I_3 - I_1)\Omega_3^2 - \ell^2 + 2I_1E}{I_2(I_1 - I_2)}, \tag{3.41b}$$

from which we get that Euler equation along axis 3 is solvable in quadratures as

$$\begin{aligned} &I_3\dot{\Omega}_3 + \Omega_1(\Omega_3)\Omega_2(\Omega_3)(I_2 - I_1) = 0, \\ \Rightarrow \quad &\int \frac{d\Omega_3}{\Omega_1(\Omega_3)\Omega_2(\Omega_3)} = \int\left(\frac{I_1 - I_2}{I_3}\right)dt. \end{aligned} \tag{3.42}$$

Writing this out explicitly gives the quadrature

$$t = \frac{I_3}{I_1 - I_2} \int \frac{d\Omega_3 \sqrt{I_1 I_2 (I_1 - I_2)^2}}{\sqrt{[I_3(I_2 - I_3)\Omega_3^2 + \ell^2 - 2I_2 E]\,[I_3(I_3 - I_1)\Omega_3^2 - \ell^2 + 2I_1 E]}}, \tag{3.43}$$

for which the right-hand side is known as an *elliptic integral*. Inverting the solution to get Ω_3 as a function of time t, then gives what is known as *Jacobi elliptic functions*, which are special functions that is sometimes thought of as a generalization of periodic sine and cosine functions.

To get a geometric picture of these systems, we can go back to the conservation equations and recast them to be written as

$$\frac{L_1^2}{2I_1 E} + \frac{L_2^2}{2I_2 E} + \frac{L_3^2}{2I_3 E} = 1, \tag{3.44a}$$

$$\frac{L_1^2}{\ell^2} + \frac{L_2^2}{\ell^2} + \frac{L_3^2}{\ell^2} = 1, \tag{3.44b}$$

which are in fact equations of an elliptical (spheroid) and a spherical surface respectively. The solutions to the angular momenta in angular momentum space are then the lines traced out by the interaction of the spheroid and spherical surfaces. However, these are all defined **in** the rotating frame.

§3.2.6 Feynman's Wobbling Plate

A famous anecdote from Richard P. Fyenman was how his curiosity on wobbling plates eventually led to his Feynman diagrams and Nobel prize winning discoveries. Specifically, he was interested in the phenomena where a free falling wobbling plate with a logo on it, appeared to wobble at twice the rate of the rotating logo (this was reversed in Feynman's book due to a typo but our conclusion is supported by Prof. David Tong at Cambridge). The wobbling motion occurs due to a consistent rotation of the plate with angular momentum orthogonal to its plane, while experiencing a back and forth oscillatory motion in the transverse direction (visualization in Fig. 3.3).

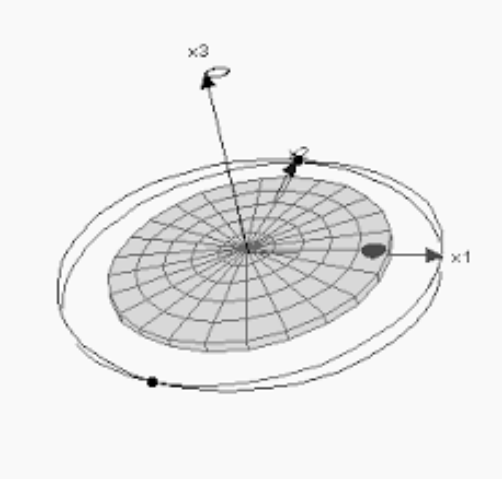

Figure 3.3: Visualization of the Feynman wobbling plate problem.

This problem is just that of a free-falling symmetric top, where the wobbling occurs due to a precession of axis 3 (symmetry axis of the plate) about the vertical lab-frame axis, z. Recall that the precession frequency is given by

$$\Omega_p = \frac{\ell}{I_1}, \tag{3.45}$$

and the rotation frequency of the plate about its axis of symmetry is given as

$$\Omega_3 = \frac{\ell \cos\theta}{I_3}, \tag{3.46}$$

where ℓ is the magnitude of the conserved angular momentum point in the z-direction. For flat objects, we have that $I_3 = I_1 + I_2 = 2I_1$, so this gives us that

$$\Omega_3 = \frac{\ell \cos\theta}{2I_1}, \tag{3.47}$$

for which in the small-angle approximation limit ($\cos\theta \approx 1$), we have

$$\Omega_3 \approx \frac{\ell}{2I_1} = \frac{\Omega_p}{2}, \tag{3.48}$$

which indeed shows that the precession, and thus wobble frequency is twice that of the rotation frequency of the logo.

As a quick recap, let us revisit all the major results we got from the study of rigid body motion. First, we found that the kinetic energy of a rigid body can be written as

$$T = \frac{1}{2}M\boldsymbol{V}^2 + \boldsymbol{\Omega}\cdot\boldsymbol{L}_{\mathrm{CoM}} + \frac{1}{2}\boldsymbol{\Omega}^T\boldsymbol{I}\boldsymbol{\Omega}. \tag{3.49}$$

To work in the body-frame of the rotating rigid body, we can work in terms of Euler angles θ, ϕ and ψ, which grants the kinetic energy

$$T = \frac{1}{2}\left[I_1\left(\dot{\phi}\sin\theta\sin\psi + \dot{\theta}\cos\psi\right)^2 + I_2\left(\dot{\phi}\sin\theta\cos\psi - \dot{\theta}\sin\psi\right)^2 + I_3\left(\dot{\phi}\cos\theta + \dot{\psi}\right)^2\right]. \tag{3.50}$$

We also derived the Euler equations, which are a general set of equations that have the angular velocities as dynamical variables and account for external torques on the system as well:

$$I_1\dot{\Omega}_1 + \Omega_2\Omega_3(I_3 - I_2) = \sum_k \tau_1^k, \tag{3.51a}$$

$$I_2\dot{\Omega}_2 + \Omega_3\Omega_1(I_1 - I_3) = \sum_k \tau_2^k, \tag{3.51b}$$

$$I_3\dot{\Omega}_3 + \Omega_1\Omega_2(I_2 - I_1) = \sum_k \tau_3^k. \tag{3.51c}$$

§3.3 Motion in Non-Inertial Frames

Having been exposed a little to considering various frames (e.g. the body-frame vs the lab-frame) in a singular system, we are now going to extend this idea further to *non-inertial* reference frames. When we study dynamics in a non-inertial (accelerating) reference frame, we will observe what are known as *fictitious forces* arise. to start, we consider 2 reference frames (primed and unprimed) which move relative to each other. One of these frames is stationary, which we denote as unprimed, while the other moves relation to this frame which we denote with primes. The velocity of an object in the primed frame as viewed from an observed in the unprimed frame can then be written as

$$\boldsymbol{v} = \boldsymbol{V}(t) + \boldsymbol{v}', \tag{3.52}$$

where $\boldsymbol{v}'$ is the velocity of the object in the primed frame, and $\boldsymbol{V}(t)$ the velocity of the primed frame relative to the unprimed frame. Considering the action of this free particle, we have

$$\begin{aligned} S &= \int dt \frac{m\boldsymbol{v}^2}{2} \\ &= \int dt \left[\frac{m\boldsymbol{V}^2}{2} + \frac{m(\boldsymbol{v}')^2}{2} + m\boldsymbol{V}\cdot\boldsymbol{v}' \right]. \end{aligned} \tag{3.53}$$

Because $\boldsymbol{V}(t)$ is just a pure function of time, it drops out of the action which leaves us with

$$\begin{aligned} S &= \int dt \left[\frac{m(\boldsymbol{v}')^2}{2} + m\boldsymbol{V}\cdot\boldsymbol{v}' \right] \\ &= \int dt \frac{m(\boldsymbol{v}')^2}{2} + \int dt \frac{d}{dt}\left(m\boldsymbol{r}'\cdot\boldsymbol{V}\right) - \int dt\, m\boldsymbol{r}'\cdot\frac{d\boldsymbol{V}}{dt}. \end{aligned} \tag{3.54}$$

This thus results in the equation of motion

$$m\dot{\boldsymbol{v}}' = -m\frac{d\boldsymbol{V}}{dt}, \tag{3.55}$$

where $d\boldsymbol{V}/dt$ is the acceleration of the reference frame. So we see that the sheer presence of a non-inertial reference frame introduce a force-like term into the equation of motion. In general, for a particle in the presence of other forces as well, we have

$$\boxed{m\dot{\boldsymbol{v}}' = \sum_j \boldsymbol{F}_j - m\frac{d\boldsymbol{V}}{dt}}. \tag{3.56}$$

Note: Einstein extended this idea in his general theory of relativiy, where it would be impossible to differentiate the force experienced due to an accelerating frame and that due to the presence of a uniform gravitational field. He termed this the *principle of equivalence*.

An example of such a fictitious force which we can experience in our daily lives is being in an elevator. When an elevator accelerates upward, it feels as though we are being pushed harder

into the ground (much like the force of gravity has increased). However, this is definitely not the case and the real physical force is simply the force from the elevator that requires to accelerating us upward.

Apart from the linear acceleration, we could also consider non-inertial reference frames which rotates (i.e. has an angular momentum that varies in time but **no** linear acceleration). To analyze such systems, we recall that in a rotating frame, the velocity in the primed reference frame would be given by

$$\boldsymbol{v}' = \boldsymbol{v} + \boldsymbol{\Omega} \times \boldsymbol{r}, \tag{3.57}$$

where $\boldsymbol{r}$ is the position of the particle relative to the unprimed frame. Plugging this into the action for a free particle gives

$$\begin{aligned} S &= \frac{m}{2} \int dt \, (\boldsymbol{v} + \boldsymbol{\Omega} \times \boldsymbol{r})^2 \\ &= \frac{m}{2} \int dt \left[\boldsymbol{v}^2 + (\boldsymbol{\Omega} \times \boldsymbol{r})^2 + 2\boldsymbol{v} \cdot (\boldsymbol{\Omega} \times \boldsymbol{r}) \right]. \end{aligned} \tag{3.58}$$

To simplify this, we use the vector identity product

$$(\boldsymbol{A} \times \boldsymbol{B}) \cdot (\boldsymbol{C} \times \boldsymbol{D}) = (\boldsymbol{A} \cdot \boldsymbol{C}) (\boldsymbol{B} \cdot \boldsymbol{D}) - (\boldsymbol{A} \cdot \boldsymbol{D}) (\boldsymbol{B} \cdot \boldsymbol{C}), \tag{3.59}$$

which results in the equations of motion (Euler-Lagrange equations)

$$\frac{\partial \mathcal{L}}{\partial \dot{\boldsymbol{r}}} = m \left[\dot{\boldsymbol{r}} + (\boldsymbol{\Omega} \times \boldsymbol{r}) \right] \tag{3.60}$$

$$\frac{\partial \mathcal{L}}{\partial \boldsymbol{r}} = m \left[\dot{\boldsymbol{r}} \times \boldsymbol{\Omega} + \boldsymbol{\Omega} \times (\boldsymbol{r} \times \boldsymbol{\Omega}) \right] \tag{3.61}$$

$$\Rightarrow \quad \boxed{m\ddot{\boldsymbol{r}} = -m \left(\dot{\boldsymbol{\Omega}} \times \boldsymbol{r} \right) + 2m (\dot{\boldsymbol{r}} \times \boldsymbol{\Omega}) + m\boldsymbol{\Omega} \times (\boldsymbol{r} \times \boldsymbol{\Omega})}, \tag{3.62}$$

where the second term on the right-hand side is known as the *Coriolis force* $\boldsymbol{F}_C$, while the third term on the right-hand side is known as the *centrifugal force* $\boldsymbol{F}_c$.

Note: The Coriolis force gives rise to motion in Foucault's pendulum, which is due to and gives a proof of the rotation of the Earth.

Perhaps a more intuitive way to write the centrifugal force would be as

$$\begin{aligned} \boldsymbol{F}_c &= m\boldsymbol{\Omega} \times (\boldsymbol{r} \times \boldsymbol{\Omega}) \\ &= m \left[\boldsymbol{r} |\boldsymbol{\Omega}|^2 - \boldsymbol{\Omega} (\boldsymbol{\Omega} \cdot \boldsymbol{r}) \right], \end{aligned} \tag{3.63}$$

where we can see that it arises as the difference of 2 terms, one along the direction of the mass' position and the other along the axis of rotation.

Chapter 4

Oscillations

One can argue that the study of oscillations is the most fundamental subject in physics, as it extends conceptually into every area and discipline of our modern theories. Furthermore, its applications are extremely relevant to engineer, rocket science and many other related fields to technology.

§4.1 1D Single Particle Oscillations

To begin our study of oscillations, the simplest system to consider is a single particle which oscillates in one-dimension close to a potential minima. The action for such a particle is given by

$$S = \int dt \left[\frac{m\dot{x}^2}{2} - V(x) \right], \tag{4.1}$$

for which if we Taylor expand the potential energy about its minimum, we get

$$V(x) \approx V(x_0) + \frac{1}{2}(x - x_0)^2 V''(x_0), \tag{4.2}$$

which is known as a *harmonic potential*. Putting this into the action gives

$$S = \int dt \left[\frac{1}{2}m\dot{x}^2 - \frac{1}{2}k(x - x_0)^2 \right], \tag{4.3}$$

where we define $k \equiv V''(x_0)$ which is always a positive quantity (by definition of a minima) known as *Hooke's constant*. Without loss of generality, we can always set-up the coordinate system such that $x_0 = 0$, so we have the more concise action

$$S = \int dt \left[\frac{1}{2}m\dot{x}^2 - \frac{1}{2}kx^2 \right], \tag{4.4}$$

$$\Rightarrow \quad m\ddot{x} = -kx, \tag{4.5}$$

$$\Rightarrow \quad x(t) = A\cos(\omega t + \varphi) = \mathrm{Re}\left\{ Ae^{i\varphi}e^{i\omega t} \right\}, \tag{4.6}$$

where $\omega \equiv \sqrt{k/m}$ and A, φ are constants defined by the initial conditions of the system. From this, we can compute the conserved total energy as

$$\begin{aligned} E &= \frac{m\dot{x}^2}{2} + \frac{kx^2}{2} \\ &= \frac{m\omega^2}{2} A^2 \sin^2(\omega t + \varphi) + \frac{k}{2} A^2 \cos^2 \omega t + \varphi \\ &= \frac{k}{2} A^2. \end{aligned} \tag{4.7}$$

§4.1.1 Forced Oscillators and Resonance

We are now going to consider a single particle subject to a potential with time-independent and time-dependent components, i.e.:

$$V(x) = V_0(x) + V_{\text{ext}}(x, t). \tag{4.8}$$

Taylor expanding around a minimum set at $x_0 = 0$, we get

$$V(x) \approx V_0(x) + V_{\text{ext}}(0, t) + V_0''(0) + xV_{\text{ext}}'(0, t), \tag{4.9}$$

which when put back into the action, gives

$$\begin{aligned} S &= \int dt \left[\frac{1}{2} m\dot{x}^2 - \frac{1}{2} kx^2 + xV_{\text{ext}}'(0, t) \right] \\ \Rightarrow \quad \ddot{x} + \omega^2 x &= \frac{F(t)}{m}, \end{aligned} \tag{4.10}$$

where $F(t) = V_{\text{ext}}'(0, t)$ is the external force acting on the system. Let's then consider a specific form of this driving time-dependent force:

$$F(t) = F_0 \cos(\gamma t + \beta). \tag{4.11}$$

From the study of differential equations, we know that with this external force, the solution of this equation of motion (by adding homogeneous and particular solutions) can be written as

$$x(t) = A \cos(\omega t + \varphi) + \frac{F_0}{m(\omega^2 - \gamma^2)} \cos(\gamma t + \beta). \tag{4.12}$$

The form of the particular solution (second term) which comes from the external force term appears to diverge if $\omega = \gamma$. We call this regime *resonance*, which corresponds to the case where the periodic driving on an oscillator matches the natural frequency of oscillation. In the regime of a resonance ($\omega = \gamma$), we can write the solution in another form (which is another valid solution) which is indicative of the physical behavior, and that is

$$x_{\text{res}}(t) = A \cos(\omega t + \varphi) + \frac{F_0 t}{2m\omega} \sin(\omega t + \beta), \tag{4.13}$$

which clearly shows a linear growth of the amplitude with time.

§4.1.2 Damped Oscillations

To extend the discussion of oscillations, we also need to consider system which undergo dissipative forces such as viscous drag. This can be modelled as a force which is proportional to the velocity of the particle such that the equation of motion of the (undriven) oscillator becomes

$$\ddot{x} + 2\lambda\dot{x} + \omega_0^2 x = 0, \tag{4.14}$$

where λ is a *damping coefficient* and the factor 2 is inserted for mathematical convenience. From this, we use the ansatz $x(t) = e^{i\Omega t}$ which when plugged back into the equation of motion, gives

$$\Omega^2 - 2i\lambda\Omega - \omega_0^2 = 0 \tag{4.15}$$

$$\Rightarrow \quad \Omega_\pm = i\lambda \pm \sqrt{\omega_0^2 - \lambda^2} \tag{4.16}$$

$$\Rightarrow \quad \boxed{x(t) = \mathrm{Re}\left[Ce^{-\lambda t}e^{\pm i\sqrt{\omega_0^2-\lambda^2}t}\right]}. \tag{4.17}$$

So we see that the solutions to the damped oscillator decay via a decay rate λ, and oscillates with frequency $\omega = \sqrt{\omega_0^2 - \lambda^2}$. There are however various regimes for a damped oscillator, for which we list these below.

1. $\underline{\lambda = 0}$:

   Here, we see that our solution parameter becomes $\Omega_\pm = \pm i\omega_0$. With this, we simply retrieve the result previously attained for the case of no damping. This case is known as **no damping**.

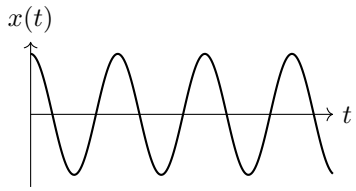

Figure 4.1: No Damping

2. $\underline{\lambda^2 > \omega_0^2}$:

   With this, we get that both λ_1 and λ_2 are real numbers. As such, we see that the system will **only** decay exponentially. In this case, the system is said to be **over damped**.

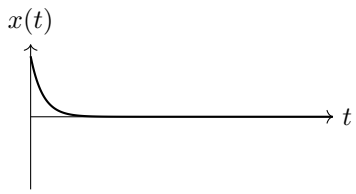

Figure 4.2: Over Damping

3. $\lambda^2 < \omega_0^2$:

   With this, we get that both λ_1 and λ_2 are complex numbers. As such, we get oscillations restricted to an exponential decay envelope. In this case, the system is said to be **under damped**. The frequency of this system deviates from the natural undamped frequency ω_0 by $\omega = \sqrt{\omega_0^2 - \lambda^2}$.

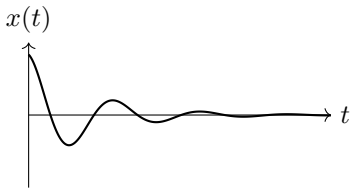

Figure 4.3: Under Damping

4. $\lambda^2 = \omega_0^2$:

   For this, we also get a pure exponential decay as we only have a $-\lambda$ as a prefactor in the exponent. In this case, the system is said to be **critically damped**.

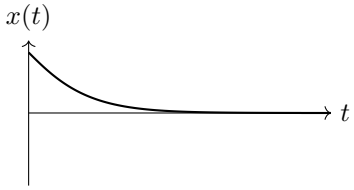

Figure 4.4: Critical Damping

   This solution, though similar to the over damped case, differs from that since we have a linear factor in t attached to the exponent as opposed to purely exponential decay terms:

$$x(t) = e^{-\lambda t}(A + Bt). \tag{4.18}$$

We now continue to augment our system for further analysis by adding an external time-dependent force in addition to damping. An example for such a system is provided in Fig. 4.5 below.

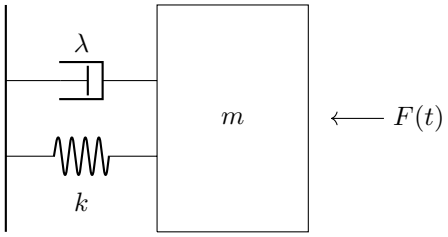

Figure 4.5: Spring-damper-mass system.

For simplicity, we can consider an external periodic force that is a pure sinusoidal function:

$$F(t) = F_0 \sin(\omega t). \tag{4.19}$$

Note that the ω used for the external force here is **not** the characteristic frequency of our undriven system. As such, our equation of motion can be written as:

$$\ddot{x} + 2\lambda\dot{x} + \frac{k}{m}x = \frac{F_0}{m}\sin(\omega t), \tag{4.20}$$

where we defined λ in this way for mathematical convenience. We learn in our study of ordinary differential equations that a general solution to a *non-homogeneous, linear differential equation* can be constructed by taking a linear combination of the *homogeneous solution* and a *particular solution*. As such, we get that our solution to this equation of motion is written as:

$$\begin{aligned} x(t) &= x_h(t) + x_p(t) \\ &= \left[C_1 e^{\left(-\lambda+\sqrt{\lambda^2-\omega_0^2}\right)t} + C_2 e^{\left(-\lambda-\sqrt{\lambda^2-\omega_0^2}\right)t}\right] + \left[A\sin(\omega t) + B\cos(\omega t)\right]. \end{aligned} \tag{4.21}$$

The thought process used to find our particular solution is as follows. First, our particular solution must satisfy the follow relation:

$$\ddot{x}_p(t) + 2\lambda\dot{x}_p(t) + \omega_0^2 x_p(t) = \frac{F_0}{m}\sin(\omega t). \tag{4.22}$$

Physically, we know that because the external driving does not die out even if the parameters would ordinarily cause the natural oscillations of the undriven system to do so, we conclude that there must be consistent oscillations occurring in the long run (steady state). This must be encoded in the particular solution and as such, we can construct the particular solution as a linear combination of sine and cosine functions as done above. Looking just at the differential equation for $x_p(t)$:

$$\begin{aligned} & (\omega_0^2 - \omega^2)\left[A\sin(\omega t) + B\cos(\omega t)\right] + 2\lambda\omega\left[A\cos(\omega t) - B\sin(\omega t)\right] = \frac{F_0}{m}\sin(\omega t), \\ \Rightarrow \quad & \left[(\omega_0^2 - \omega^2)A - 2\lambda\omega B\right]\sin(\omega t) + \left[(\omega_0^2 - \omega^2)B + 2\lambda\omega A\right]\cos(\omega t) = \frac{F_0}{m}\sin(\omega t), \\ \Rightarrow \quad & (\omega_0^2 - \omega^2)A - 2\lambda\omega B = \frac{F_0}{m}, \quad (\omega_0^2 - \omega^2)B + 2\lambda\omega A = 0. \end{aligned} \tag{4.23}$$

Note: We could look at the sine and cosine prefactors independent since sine and cosine are *orthogonal functions*. That is to say:

$$\frac{1}{T}\int_0^T \sin(\omega t)\cos(\omega t)dt = 0 \tag{4.24}$$

where $T = \frac{2\pi}{\omega}$ is the driving period (the integral over one period is the *inner product* for continuous, periodic functions).

As such, we can write the 2 equations for A and B as a linear system in matrix form as follows:

$$\begin{bmatrix} \omega_0^2 - \omega^2 & -2\lambda\omega \\ 2\lambda\omega & \omega_0^2 - \omega^2 \end{bmatrix} \begin{bmatrix} A \\ B \end{bmatrix} = \begin{bmatrix} \frac{F_0}{m} \\ 0 \end{bmatrix}, \tag{4.25}$$

$$\Rightarrow \quad A = \frac{\frac{F_0}{m}(\omega_0^2 - \omega^2)}{(\omega_0^2 - \omega^2)^2 + (2\lambda\omega)^2}, \quad B = \frac{-2\lambda\omega\frac{F_0}{m}}{(\omega_0^2 - \omega^2)^2 + (2\lambda\omega)^2}, \tag{4.26}$$

$$\Rightarrow \quad x_p(t) = \frac{\left(\frac{F_0}{m}\right)}{(\omega_0^2 - \omega^2)^2 + (2\lambda\omega)^2} \left[(\omega_0^2 - \omega^2)\sin(\omega t) - 2\lambda\omega\cos(\omega t)\right]. \tag{4.27}$$

To extract more physical insights to this result, we recall the trigonometric identity $A\sin(x) + B\cos(x) = \sqrt{A^2 + B^2}\sin(x + \phi)$ where ϕ is a phase given by $\tan^{-1}(-B/A)$. Utilizing this, we get that our particular solution becomes:

$$x_p(t) = \frac{\left(\frac{F_0}{m}\right)}{\sqrt{(\omega_0^2 - \omega^2)^2 + (2\lambda\omega)^2}} \sin\left[\omega t + \tan^{-1}\left(\frac{2\lambda\omega}{\omega_0^2 - \omega^2}\right)\right]. \tag{4.28}$$

So our general solution for this system can be written as:

$$\begin{aligned} x(t) = &\left[C_1 e^{\left(-\lambda + \sqrt{\lambda^2 - \omega_0^2}\right)t} + C_2 e^{\left(-\lambda - \sqrt{\lambda^2 - \omega_0^2}\right)t}\right] \\ &+ \frac{\left(\frac{F_0}{m}\right)}{\sqrt{(\omega_0^2 - \omega^2)^2 + (2\lambda\omega)^2}} \sin\left[\omega t + \tan^{-1}\left(\frac{2\lambda\omega}{\omega_0^2 - \omega^2}\right)\right]. \end{aligned} \tag{4.29}$$

As further analysis, we can generate a plot of the amplitude of $x_p(t)$ against ω (as shown in Fig. 4.6 below) to see how the steady state solution amplitude varies with ω.

Note: We often refer to the particular solution as the steady state solution since we have that the homogeneous solution will be damped out after a long time (decays away), leaving only the particular solution.

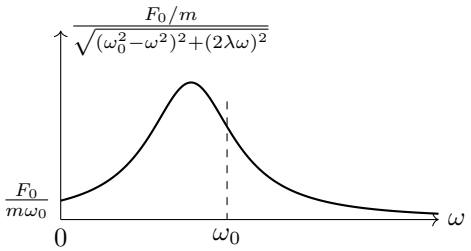

Figure 4.6: Oscillation Amplitude vs Angular Frequency Plot

When divided by the amplitude of the drive and its frequency squared $\omega^2 F_0/m$, the function in figure 4.6 becomes what is known as the *transmissibility*, τ. The transmissiblity of a system is a unit-free measure of its response to driving, more generally defined as $\tau = \text{output/input}$. There are several things we can extract from the resonance plot above:

- We see that if $\omega >> \omega_0$, we get that $\frac{\left(\frac{F_0}{m}\right)}{\sqrt{(\omega_0^2-\omega^2)^2+(2\lambda\omega)^2}} \to 0$.
- We see that the maximum amplitude does **not** occur at ω_0, although close by. The maximum amplitude only coincides with ω_0 when $\lambda = 0$.

One can now ask about the energy dissipated from the system due to the damping. To determine this, we can multiply the damped oscillator equation of motion by $\dot{x}$ which gives

$$\dot{x}\ddot{x} + 2\lambda\dot{x}^2 + \omega^2 x\dot{x} = \frac{F\dot{x}}{m}. \tag{4.30}$$

We can see that the term on the right hand side of the equation above is in fact the power (divided by m) $P_{\text{in}} = F\dot{x}$, being pumped into the system. This leaves us with the equation

$$\frac{d}{dt}\left(\frac{m\dot{x}^2}{2} + \frac{m\omega^2 x^2}{2}\right) + 2\lambda m\dot{x}^2 = F\dot{x}. \tag{4.31}$$

From this, we see that the power loss from the system due to damping is given by

$$\boxed{P_{\text{loss}} = -2\lambda m\dot{x}^2}. \tag{4.32}$$

The energy dissipated is then the integral over time of the power, given as

$$\boxed{E_{\text{loss}} = -2\lambda m\int_{t_i}^{t_f} \dot{x}^2(t)dt}. \tag{4.33}$$

§4.2 Multi-Mode Oscillators

In this section, we will be looking at oscillators of multiple modes (a system of several coupled oscillators). As such, these would turn out to be systems in which the energies are functions of several generalized coordinates ($T = T(\dot{q}_1, \dot{q}_2, ..., \dot{q}_n)$ and $V = V(q_1, q_2, ..., q_n)$). For a purely mechanical system, we have that the kinetic energy is given by:

$$T = \sum_j \frac{1}{2} m_j \dot{q}_j^2 \tag{4.34}$$

Now if we consider a region close to the equilibrium position of the system and only small perturbations around this position, we have that:

$$V \approx \sum_{ij} \frac{1}{2} k_{ij} q_i q_j \tag{4.35}$$

This is because, performing a Taylor expansion around any local potential-well gives us a quadratic potential (ignoring higher order terms and taking that at equilibrium, $\boldsymbol{\nabla} V(q_j^*) = 0$). Knowing Lagrangian mechanics, we can write down the equations of motion from these energy functions as:

$$\ddot{q}_i + \frac{k_{ii}}{m_i} q_i + \sum_{j\neq i} \frac{k_{ij}}{2m_i} q_j = 0 \tag{4.36}$$

where the term that sums over index j represents the couplings between adjacent oscillators in the oscillator lattice. If we take the ansatz to be $q_i(t) = C_i e^{i\omega t}$, this gives us:

$$\omega^2 q_i = \frac{k_{ii}}{m_i} q_i + \sum_{j \neq i} \frac{k_{ij}}{2m_i} q_j \tag{4.37}$$

Note: Be sure to distinguish between the index i and complex number $i = \sqrt{-1}$. This is somewhat sloppy notation but relatively unambiguous.

We can then represent these equations indexed by i as a linear system in matrix form as follows:

$$\begin{bmatrix} \ddots & & & & \\ & & \frac{k_{i-1,i+1}}{2m_i} & & \\ & \frac{k_{i,i}}{m_i} & & & \\ & \frac{k_{i+1,i-1}}{2m_i} & & & \\ & & & & \ddots \end{bmatrix} \begin{bmatrix} \vdots \\ q_{i-1} \\ q_i \\ q_{i+1} \\ \vdots \end{bmatrix} = \omega^2 \begin{bmatrix} \vdots \\ q_{i-1} \\ q_i \\ q_{i+1} \\ \vdots \end{bmatrix} \tag{4.38}$$

We actually also have that $\frac{k_{i,j}}{2m_i} = \frac{k_{j,i}}{2m_i}$, meaning that the matrix of spring constants and reciprocal masses is in fact **symmetric** (due to Newton's second law). We can relabel the linear system above in the following way:

$$(M^{-1}K)\boldsymbol{q} = \omega^2 \boldsymbol{q} \tag{4.39}$$

where the matrix entries M_{ij} indicate adjacent masses whereas the entries K_{ij} indicate the spring constants connecting masses m_i and m_j. We see that in fact, the equation above is an *eigenvalue* problem, where we require to look for the eigen-decomposition of the matrix $M^{-1}K$ (with eigenvalues ω_i^2 and eigenvectors $\boldsymbol{q}$). We often refer to the eigenvalues of the $M^{-1}K$ matrix as the *eigen-frequencies*, and the eigenvectors as *eigen-modes*. By closely following the chosen ansatz, we can already draw several interesting conclusions by simply looking at the parity of ω^2.

1. $\underline{\omega^2 > 0}$:

   Here, we get that q_i is a linear combination of sines and cosines, causing the system to undergo oscillations indefinitely. The functional form would thus be:

   $$q(t) = Ae^{i\omega t} + Be^{-i\omega t} \tag{4.40}$$

2. $\underline{\omega^2 < 0}$:

   Here, we get that $q_i \propto e^{i(\pm i\omega)t}$, which means we get a linear combination of real exponential functions:

   $$q(t) = Ae^{\omega t} + Be^{-\omega t} \tag{4.41}$$

If we further assert that the initial conditions of our system are $q(0) = \varepsilon$ and $\dot{q}(0) = 0$, then we arrive at the solution:

$$q(t) = \frac{\varepsilon}{2}(e^{\omega t} + e^{-\omega t}) = \varepsilon \sinh(\omega t) \tag{4.42}$$

which means that the generalized coordinate of our system grows exponentially, implying our system was likely at a maxima (or unstable point) to begin with.

Let's now look at an example to learn how to utilize this method of eigen-modes and eigen-frequencies to solve an oscillating system.

Example:

Consider the following coupled system of springs and masses as seen in figure 4.7.

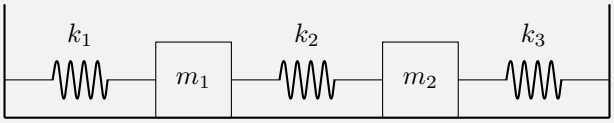

Figure 4.7: Multi-mode system of 3 springs and 2 masses.

With this, we can write the equations of motion as:

$$m_1\ddot{x}_1 = -k_1x_1 + k_2(x_2 - x_1) \tag{4.43}$$
$$m_2\ddot{x}_2 = -k_2(x_2 - x_1) + k_3x_2 \tag{4.44}$$

where x_1 and x_2 denote the displacements of masses m_1 and m_2 respectively. We write this in a matrix representation as follows:

$$\begin{bmatrix} m_1 & 0 \\ 0 & m_2 \end{bmatrix} \begin{bmatrix} \ddot{x}_1 \\ \ddot{x}_2 \end{bmatrix} = \begin{bmatrix} -k_1 - k_2 & k_2 \\ k_2 & -k_2 - k_3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \tag{4.45}$$

Now we pick the following ansatz and assert the following definitions:

$$\boldsymbol{x} = e^{i\omega t}\boldsymbol{x}_0 \tag{4.46}$$
$$M = \begin{bmatrix} m_1 & 0 \\ 0 & m_2 \end{bmatrix}, \quad K = \begin{bmatrix} k_1 + k_2 & -k_2 \\ -k_2 & k_2 + k_3 \end{bmatrix} \tag{4.47}$$

where $\boldsymbol{x}_0$ is a time-independent eigenmode. We simply attach this complex phase ($e^{i\omega t}$) to the initial conditions as a means to find the ω solutions. We also know that these complex exponentials give general solutions to second order differential equations. Plugging in this ansatz produces:

$$\begin{aligned} & \omega^2 M\boldsymbol{x}_0 = K\boldsymbol{x}_0 \\ \Rightarrow \quad & M^{-1}K\boldsymbol{x}_0 = \omega^2\boldsymbol{x}_0 \end{aligned} \tag{4.48}$$

For further simplicity, let us assume that $m_1 = m_2 = m$ and $k_1 = k_2 = k_3 = k$. With

this, we can solve the eigenvalue problem as follows:

$$\begin{aligned} & \det\left(M^{-1}K - \mathbb{I}\omega^2\right) = 0 \\ \Rightarrow\ & \det\left\{\frac{k}{m}\begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix} - \omega^2 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\right\} = 0 \\ \Rightarrow\ & (2 - \frac{m\omega^2}{k})^2 - 1 = 0 \\ \Rightarrow\ & \frac{m\omega^2}{k} = 1 \ \text{ or } \ 3 \end{aligned} \tag{4.49}$$

$$\Rightarrow\ \omega = \sqrt{\frac{3k}{m}} \ \text{ or } \ \sqrt{\frac{k}{m}} \tag{4.50}$$

As for the eigen-modes, we have to find the eigen-modes corresponding to each eigen-frequency found above as follows:

1. $(\frac{m\omega^2}{k} = 1)$:

$$\begin{aligned} & \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} - \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \mathbf{0} \\ \Rightarrow\ & x_1 = x_2 \\ \Rightarrow\ & \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \end{aligned} \tag{4.51}$$

   This eigen-mode is telling us is that the 2 masses are moving by the same displacement in the same direction 'in sync'.

2. $(\frac{m\omega^2}{k} = 3)$:

$$\begin{aligned} & \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} - 3\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \mathbf{0} \\ \Rightarrow\ & x_1 = -x_2 \\ \Rightarrow\ & \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \end{bmatrix} \end{aligned} \tag{4.52}$$

   This eigen-mode is telling us is that the 2 masses are moving by the same displacement in opposite directions.

Realize that from the example above, we see that the 2 eigen-modes are *orthogonal*. In fact, this is true for all eigen-modes since the $M^{-1}K$ matrix will always be a real symmetric matrix (by the *spectral theorem*). Moreover, these eigen-modes *span* the **entire** vector space of possible vibrational modes! As such, we can always write any arbitrary vibration of a system as a linear combination of its eigen-modes with their corresponding phases tacked on. This is more clearly written as:

$$\boxed{\boldsymbol{x}(t) = \sum_j A_j e^{i\omega_j t}\boldsymbol{\eta}_j} \tag{4.53}$$

where A_j are real coefficients and $\boldsymbol{\eta}_j$ are the eigen-modes with associated eigen-frequencies ω_j.

Chapter 5

Hamiltonian Formalism

Having studied the dynamics of mechanical system for the past several chapters, it is appropriate now to shift gears and look at the equations of motion themselves in more detail. In particular, we want to consider the Hamiltonian picture of classical mechanics, which leads to Hamilton's equations of motion and a geometric picture of mechanical systems via the language of phase space.

§5.1 Hamilton's Equations of Motion

Thus far, we have utilized the principle of least action with a Lagrangian function to derive the equations of motion, dubbed the Euler-Lagrange equations. An alternative means to derive equations of motion is to start with another function known as the Hamiltonian. The Hamiltonian is a function of the generalized coordinates q_j, and generalized momenta

$$p_j = \frac{\partial \mathcal{L}}{\partial \dot{q}_j}. \tag{5.1}$$

That is, the move from a Lagrangian to a Hamiltonian trades the use of $\dot{q}_j$ for the use of p_j instead. The Hamiltonian written in terms of the Lagrangian is written as

$$\boxed{\mathcal{H}(q_j, p_j) = \sum_j p_j \dot{q}_j - \mathcal{L}}, \tag{5.2}$$

which is motivated by the total energy of the system (we take that $\dot{q}_j = \dot{q}_j(p_j)$). A generalization of the transformation above (from $\mathcal{L}$ to $\mathcal{H}$) is known as a *Legendre transformation*, in which for any function $f(x)$, its Legendre transform is given by

$$g(y) = x\frac{\partial f(x)}{\partial x} - f(x) = xy(x) - f(x), \tag{5.3}$$

where $y = \partial f/\partial x$. To put simply, this is a way to trade one function for another and has applications also in statistical physics (e.g. to move between the various forms of free energy).

Notice that there is a symmetry to this transformation, in that

$$y(x) = \frac{\partial f(x)}{\partial x}, \tag{5.4a}$$

$$x(y) = \frac{\partial g(y)}{\partial y}. \tag{5.4b}$$

So going back to Hamiltonian mechanics, we see that $\mathcal{H}$ is just the Legendre transform of $\mathcal{L}$ with respect to velocities. The first dynamical equation we can extract from $\mathcal{H}$ is simply that

$$\dot{q}_j = \frac{\partial \mathcal{H}}{\partial p_j}. \tag{5.5}$$

This however, would not give a comprehensive picture of the dynamics we require an equation for the time derivative of momenta. To obtain this, first consider

$$\begin{aligned} \frac{\partial \mathcal{H}}{\partial q_k} &= \sum_j p_j \frac{\partial \dot{q}_j}{\partial q_k} - \frac{\partial \mathcal{L}}{\partial q_k} - \sum_j \frac{\partial \mathcal{L}}{\partial \dot{q}_j}\frac{\partial \dot{q}_j}{\partial q_k} \\ &= -\frac{\partial \mathcal{L}}{\partial q_k} \\ &= -\frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{q}_k} = \dot{p}_k, \end{aligned} \tag{5.6}$$

where we utilized the Euler-Lagrange equation above. As a result, we have that

$$\boxed{\dot{p}_k = -\frac{\partial \mathcal{H}}{\partial q_k}, \quad \dot{q}_j = \frac{\partial \mathcal{H}}{\partial p_j}}, \tag{5.7}$$

which are known as the *Hamilton equations of motion*. As a quick sanity check, we can consider the simple system of a single particle in 1-dimension subject to a potential $U(x)$. The associated Lagrangian and Hamiltonian for this system are then

$$\mathcal{L} = \frac{1}{2}m\dot{x}^2 - U(x), \tag{5.8a}$$

$$\mathcal{H} = p\dot{x} - \frac{1}{2}m\dot{x}^2 + U(x) = \frac{p^2}{2m} + U(x). \tag{5.8b}$$

Plugging the Hamiltonian into the Hamilton equations of motion then give the familiar results

$$\dot{p} = -\frac{\partial U(x)}{\partial x}, \tag{5.9a}$$

$$\dot{x} = \frac{p}{m}. \tag{5.9b}$$

A useful feature of using the Hamiltonian instead of the Lagrangian, is that as long as it is not an explicit function of time, the Hamiltonian will be conserved. To see this explicitly, the total time derivative of $\mathcal{H}$ is given by

$$\frac{d\mathcal{H}}{dt} = \sum_j \left(\frac{\partial \mathcal{H}}{\partial q_j}\dot{q}_j + \frac{\partial \mathcal{H}}{\partial p_j}\dot{p}_j \right) + \frac{\partial \mathcal{H}}{\partial t}. \tag{5.10}$$

So we see that if $\mathcal{H}$ is not an explicitly function of time, we have

$$\frac{d\mathcal{H}}{dt} = \sum_j \left(\frac{\partial \mathcal{H}}{\partial q_j} \dot{q}_j + \frac{\partial \mathcal{H}}{\partial p_j} \dot{p}_j \right) = \sum_j \left(\frac{\partial \mathcal{H}}{\partial q_j} \frac{\partial \mathcal{H}}{\partial p_j} - \frac{\partial \mathcal{H}}{\partial p_j} \frac{\partial \mathcal{H}}{\partial q_j} \right) = 0, \tag{5.11}$$

indeed proving Hamiltonian conservation.

§5.2 Phase Space

We now introduce the concept of phase space, which is a $2N$-dimensional space (N being the number of degrees of freedom) with axes being the coordinates $\{q_j, p_j\}$. In phase space, we can think of the Hamilton equations

$$\dot{q}_j = \frac{\partial \mathcal{H}}{\partial p_j}, \tag{5.12a}$$

$$\dot{p}_k = -\frac{\partial \mathcal{H}}{\partial q_k}, \tag{5.12b}$$

acting as generalized velocities which tells the system how to traverse the phase space. In other words, these equations define a vector field on phase space, which solutions give the trajectories

$$q_j(t) = q_j(q^0, p^0; t), \tag{5.13a}$$

$$p_j(t) = p_j(q^0, p^0; t), \tag{5.13b}$$

that in fact satisfy the group property that time translations are transitive:

$$q_j(t) = q_j[q(q^0, p^0; t_2), p(q^0, p^0; t_2); t_1], \tag{5.14a}$$

$$\text{where} \quad t = t_1 + t_2. \tag{5.14b}$$

§5.2.1 Liouville's Theorem

An important aspect of looking at the phase space of a *Hamiltonian system* (system that obeys Hamilton's equations of motion) is that it adheres to Liouville's theorem. The theorem said informally, tells us that the evolution of a Hamiltonian system in time does not change the amount of volume it occupies in phase space. Before getting into the formal statement of this theorem, it would be nice to gain some intuition about it. The intuition we build here stems from a statistical picture of the classical system, whereby we assume a large enough ensemble such that we can consider a coarse-graining over all particles into an effective continuous distribution function. For simplicity, we first imagine a discretized region of finite phase space elements. In this discrete space, we consider the set of all possible states $\mathcal{S}$ of a system, each state labelled by $s_k \in \mathcal{S}$. Each state then encompasses a particular configuration of the dynamical system in question.

In discrete phase space, each of these states would be comprised of a set of state constituents ("phase space chunks"), where the set of all state constituents of all possible states is defined as $\mathcal{B}$. The way these constituents amalgamate to a state is that they follow some distribution which allows us to write $s_k = \sum_j \alpha_j b_j$ where $b_j \in \mathcal{B}$ and $\alpha_j \in \mathbb{R}^+ \cup \{0\}$ such that $\sum_j \alpha_j = 1$.

An illustration of a state in discrete phase space with 10 constituents is given in Fig. 5.1. In the illustration, the group of yellow squares represent the state, each of which is a state constituent $\in \mathcal{B}$. Note that the constituents here only constitute a subset of $\mathcal{B}$. In this discrete formalism, we can treat each state as a vector in a real vector space, for which each constituent is then a basis vector. This is nice because it tells us that knowing how each basis vector evolves in time, allows us to know how **any** arbitrary state of the system evolves in time by linearity!

Now we are ready to get into state evolution in time. Classical mechanics asserts that time-evolution of a Hamiltonian system is deterministic and reversible. Mathematically, this translates to time-evolution of a state being a *bijective* (one-to-one) mapping between basis vectors (i.e. the number of basis vectors that constitute a state remains unchanged in time). As such, Hamiltonian time-evolution in phase space would look something like that shown in Fig. 5.2, where the initial state is colored yellow (lower left distribution) and the final state is colored cyan (upper right distribution). Notice how there are the same number of constituents in each state, which follow from the assertions of classical mechanics, and alludes to Liouville's theorem.

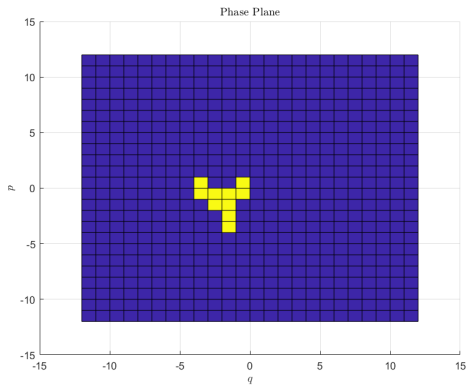

Figure 5.1: State in discrete phase space.

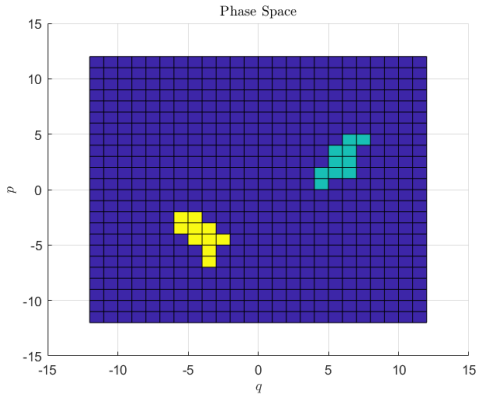

Figure 5.2: State time-evolution in phase space.

Now we get into the actual formulation of the theorem. Extension of the intuition we have built into continuous phase space requires the definition of the *phase space distribution function*.

Definition 5.2.1. *The phase space distribution function determines the probability that a system will be found in some infinitesimal phase space volume.*

From here, we state Liouville's theorem.

Theorem 5.2.1. *The phase space distribution function remains constant in Hamiltonian evolution. That is to say, the phase space hypervolume remains constant in time.*

Note that the term *hypervolume* in the theorem above is used to generalize the theorem to higher dimensions. To prove this theorem, we first need knowledge of the *divergence (Gauss') theorem*, which will be simply stated in these notes.

Theorem 5.2.2. *Given some vector field $\boldsymbol{F}$ and some region of space with volume V and boundary ∂V, the volume integral of the divergence of $\boldsymbol{F}$ and the surface integral of the field are related as follows:*

$$\int_V \boldsymbol{\nabla} \cdot \boldsymbol{F} dV = \int_{\partial V} \boldsymbol{F} \cdot d\boldsymbol{A} \tag{5.15}$$

The proof of Liouville's theorem is then given as follows:

Proof. We first prove this in the case of a single degree of freedom system (one pair of conjugate variables). Let us first think of the trajectory of some differential area of phase space:

$$\begin{bmatrix} dq \\ dp \end{bmatrix} = \begin{bmatrix} \dot{q} \\ \dot{p} \end{bmatrix} dt = \boldsymbol{v} dt \tag{5.16}$$

where $\boldsymbol{v}$ is the velocity at some instance of time over time interval dt. To compute the change in area of the whole phase space region in question over time interval dt, we perform an integral:

$$dA = \int_C dl(\boldsymbol{n} \cdot \boldsymbol{v}) dt \tag{5.17}$$

where $\boldsymbol{n}$ denotes the unit-normal vector to the boundary contour and C indicates an integral over the entire phase space boundary. From here, we can exploit the use of the *divergence theorem* to get:

$$\begin{aligned} \frac{dA}{dt} &= \int_C dl(\boldsymbol{n} \cdot \boldsymbol{v}) \\ &= \int_A \boldsymbol{\nabla} \cdot \boldsymbol{v} dA \\ &= \int_A dA \left(\frac{\partial}{\partial q} \dot{q} + \frac{\partial}{\partial p} \dot{p} \right) \\ &= \int_A dA \left(\frac{\partial}{\partial q} \frac{\partial \mathcal{H}}{\partial p} - \frac{\partial}{\partial p} \frac{\partial \mathcal{H}}{\partial q} \right) \end{aligned} \tag{5.18}$$

Since partial derivatives commute, we see that the term above vanishes $\frac{dA}{dt} = 0$ and we indeed get that the area of the phase space is conserved. The extension of this proof into systems with higher degrees of freedom is trivial since the divergence theorem is valid in N dimensions as well. □

The final result in the proof is known as *Liouville's equation*, and has the general form for a multi-degree of freedom system as presented below:

$$\boxed{\frac{d}{dt} \rho(q_j, p_j) = \frac{\partial \rho}{\partial t} + \sum_j \left(\frac{\partial \rho}{\partial q_j} \dot{q}_j + \frac{\partial \rho}{\partial p_j} \dot{p}_j \right) = 0} \tag{5.19}$$

whereas the motion of an individual member of the ensemble is given by Hamilton's equations,

Liouville's equations describe the flow of the whole distribution. Alternatively, we can prove Liouville's theorem by considering the Jacobian of the transformation in phase space. This is provided below.

Proof. Having a trajectory in phase space, we can consider the initial and final configurations of the particles denoted by the tuples (q_0, p_0) and (q, p) respectively. The transformation between these coordinates will then have the Jacobian

$$J(t) = \frac{\partial(q,p)}{\partial(q_0,p_0)}, \tag{5.20}$$

which equates to unity of Liouville's theorem holds (i.e. $\frac{d}{dt}J(t) = 0$). To ascertain this, we first note the identity

$$\frac{d}{dt}\det\{A(t)\} = \det\{A(t)\}\cdot \mathrm{Tr}\left\{A^{-1}(t)\frac{dA(t)}{dt}\right\}. \tag{5.21}$$

Now we consider the transformation occurring from time 0 to time τ to time t:

$$J(t) = \frac{\partial(q,p)}{\partial(q_\tau,p_\tau)}\cdot\frac{\partial(q,p)}{\partial(q_\tau,p_\tau)}, \tag{5.22}$$

such that we can consider the matrix associated to the Jacobian:

$$S = \begin{bmatrix} \frac{\partial q_t}{\partial q_\tau} & \frac{\partial q_t}{\partial p_\tau} \\ \frac{\partial p_t}{\partial q_\tau} & \frac{\partial p_t}{\partial p_\tau} \end{bmatrix}, \tag{5.23}$$

where the matrix is extrapolated accordingly for systems of more degrees of freedom. This matrix becomes the identity when $t = \tau$ as expected. We then consider the derivative of $\det\{S(t)\}$ with respect to t, and then take the limit where $t \to \tau$ which gives

$$\begin{aligned} \lim_{t\to\tau}\frac{d}{dt}\det\{S(t)\} &= \lim_{t\to\tau}\det\{S(t)\}\,\mathrm{Tr}\left\{S^{-1}\frac{dS}{dt}\right\} \\ &= \lim_{t\to\tau}\mathrm{Tr}\left\{\frac{dS}{dt}\right\} \\ &= \lim_{t\to\tau}\mathrm{Tr}\left\{\sum_{\#\text{ d.o.f.}}\left(\frac{\partial\dot{q}}{\partial q_\tau}+\frac{\partial\dot{p}}{\partial p_\tau}\right)\right\} \\ &= \mathrm{Tr}\left\{\sum_{\#\text{ d.o.f.}}\left(\frac{\partial^2\mathcal{H}}{\partial q\partial p}-\frac{\partial^2\mathcal{H}}{\partial p\partial q}\right)\right\} = 0. \end{aligned} \tag{5.24}$$

As such, we see that indeed

$$\lim_{t\to\tau}\frac{d}{dt}\det\{S(t)\} = \frac{d}{dt}J(t) = 0. \tag{5.25}$$

□

§5.3 Routh's Function

We are now going to introduce another formalism, which combines both the Lagrangian and Hamiltonian functions. This approach is useful in dealing with complicated systems with cyclic coordinates, which allows the cyclic coordinates to be treated in the Hamiltonian formalism and non-cyclic coordinates in the Lagrangian formalism. With this, the *Routh function* is defined as

$$\mathcal{R} = \sum_c p_c \dot{q}_c - \mathcal{L}, \tag{5.26}$$

where the sum over c runs over cyclic coordinates. With this, we obtain the equations of motion

$$\dot{q}_c = \frac{\partial \mathcal{R}}{\partial p_c}, \tag{5.27a}$$

$$\dot{p}_c = -\frac{\partial \mathcal{R}}{\partial q_c}, \tag{5.27b}$$

$$\frac{d}{dt}\frac{\partial \mathcal{R}}{\partial \dot{q}_n} = \frac{\partial \mathcal{R}}{\partial q_n}, \tag{5.27c}$$

where indices n denote non-cyclic coordinates. It would now be useful to show the usefulness of this formulation with an example.

Example:

Consider a single particle system under the influence of a central potential. The Lagrangian for this system is then written as

$$\mathcal{L}(r, \phi) = \frac{1}{2} m \dot{r}^2 + \frac{1}{2} m \dot{\phi}^2 - U(r). \tag{5.28}$$

As such, we see that ϕ is a cyclic variable for this system with 2 degrees of freedom, and we immediately have that

$$p_\phi = m r^2 \dot{\phi} = \text{constant} \equiv \ell. \tag{5.29}$$

The Routh function is then written as

$$\begin{aligned} \mathcal{R} &= \dot{\phi} p_\phi - \mathcal{L} \\ &= \ell \dot{\phi} - \mathcal{L} \\ &= \frac{\ell^2}{2mr^2} - \frac{1}{2} m \dot{r}^2 - U(r). \end{aligned} \tag{5.30}$$

The resulting equations of motion are then

$$\dot{\phi} = \frac{\partial \mathcal{R}}{\partial \ell} = \frac{\ell}{m r^2}, \tag{5.31a}$$

$$\dot{\ell} = \frac{\partial \mathcal{R}}{\partial \phi} = 0, \tag{5.31b}$$

$$\frac{d}{dt}\frac{\partial \mathcal{R}}{\partial \dot{r}} = \frac{\partial \mathcal{R}}{\partial r}, \tag{5.31c}$$

with conserved "energy" function

$$\begin{aligned}\tilde{E} &= \dot{r}\frac{\partial \mathcal{R}}{\partial \dot{r}} - \mathcal{R} \\ &= -\frac{1}{2}m\dot{r}^2 - U(r) - \frac{\ell^2}{2mr^2}.\end{aligned} \tag{5.32}$$

We see the conserved "energy" obtained here comes with an added minus sign, but is in fact easier to construct since it picks out only non-cyclic coordinates in its formulation.

§5.4 The Action: A Deep Dive

Previous treatments of the action we have come across, take it simply as a means to an end. That is, we have considered the action only as a functional to minimize which grants us the Euler-Lagrange equations of motion. However, the action itself carries valuable information when we cast it as a function of position and time. We can see this from the relations

$$\frac{\partial S}{\partial q} = p, \tag{5.33a}$$

$$\frac{\partial S}{\partial t} = -\mathcal{H}(q,p), \tag{5.33b}$$

which lead to the *Hamilton-Jacobi* equations. But before getting to that, we shall first work through the derivations to obtain the relations above. Starting with the action

$$S(q,t) = \int_{t_i}^{t} d\tau \left[\sum_j \frac{1}{2}m_j\dot{q}_j^2 - U(q_j)\right], \tag{5.34}$$

there are some general things that can be said about the action without solving for the equations of motion. First, consider a trajectory $q(\tau)$ which ends at $q(t) \equiv q$. We can then ask, what happens to the action if we perturb the endpoint by a small amount δq, such that $S = S(q+\delta q, t)$. The new trajectory is then $q(\tau) + \delta q(\tau)$, and the variation in the action is then

$$\begin{aligned}\delta S &= S(q+\delta q, t) - S(q,t) \\ &= \int_{t_i}^{t} d\tau \left[\frac{\partial \mathcal{L}}{\partial q}\delta q + \frac{\partial \mathcal{L}}{\partial \dot{q}}\delta\dot{q}\right] \\ &= \int_{t_i}^{t} d\tau \left[\frac{\partial \mathcal{L}}{\partial q} - \frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{q}}\right]\delta q + \left.\frac{\partial \mathcal{L}}{\partial \dot{q}}\delta q\right|_{t_i}^{t} \\ &= \left.\frac{\partial \mathcal{L}}{\partial \dot{q}}\delta q\right|_{t_i}^{t} = \frac{\partial \mathcal{L}}{\partial \dot{q}}(t)\delta q(t),\end{aligned} \tag{5.35}$$

since we did not vary the initial position of the trajectory. So we see that as expected, the variation in the action is directly dependent on the variation of the endpoint, $\delta q(t)$ (i.e. varying the endpoint induces a change in the entire trajectory). Rearranging the result above and

generalizing to multiple degrees of freedom then gives

$$\boxed{\frac{\partial S}{\partial q_j} = \frac{\partial \mathcal{L}}{\partial \dot{q}_j} = p_j}, \tag{5.36}$$

which is the first expression we wanted to derive.

Note: An interesting consequence of this relation is that

$$\frac{\partial^2 S}{\partial q_j \partial q_k} = \frac{\partial p_j}{\partial q_k} = \frac{\partial p_k}{\partial q_j}. \tag{5.37}$$

This is a rather peculiar result that says that somehow, the the momenta of each degree of freedom is dependent on the coordinates of all other degrees of freedom.

Now what about $\partial S/\partial t$? Getting this is a little more complicated since we require to consider a variation in time δt, whilst maintaining a fixed endpoint in the trajectory such that

$$q(t) = q(t + \delta t). \tag{5.38}$$

From here, we once again consider the variation in the action

$$\begin{aligned} \delta S &= S(q + \delta q, t) - S(q, t) \\ &= \int_{t_i}^{t} d\tau \left[\mathcal{L}(q + \delta q) - \mathcal{L}(q)\right] + \int_{t}^{t+\delta t} d\tau \mathcal{L}(\tilde{q}) \\ &\approx \frac{\partial \mathcal{L}}{\partial \dot{q}}(t)\delta q(t) + \mathcal{L}[q(t)]\delta t, \end{aligned} \tag{5.39}$$

where $\tilde{q}$ is the additional sliver of trajectory induced by the variation in t, and we performed a Taylor expansion keeping only the lowest order non-trivial term in δt. Next, we utilize the relation

$$\delta q(t) = -\dot{q}(t)\delta t, \tag{5.40}$$

$$\Rightarrow \quad \delta S = \left(-\frac{\partial \mathcal{L}}{\partial \dot{q}}(t)\dot{q}(t) + \mathcal{L}[q(t)]\right)\delta t = -\mathcal{H}\delta t. \tag{5.41}$$

Rearranging this and generalizing, we indeed get

$$\boxed{\frac{\partial S}{\partial t} = -\frac{\partial \mathcal{L}}{\partial \dot{q}_j}\dot{q}_j + \mathcal{L}(q_j, t) = -\mathcal{H}}. \tag{5.42}$$

From this, we can get the Hamilton's equation of motion if we consider

$$\frac{\partial^2 S}{\partial t \partial q_j} = -\frac{\partial \mathcal{H}}{\partial q_j} = \frac{\partial p_j}{\partial t}. \tag{5.43}$$

To go even further, we can perform a Legendre transform on the Lagrangian in the action to give

$$S = \int d\tau \mathcal{L} = \int d\tau \left[\sum_j p_j \dot{q}_j - \mathcal{H}(p,q) \right]. \tag{5.44}$$

Now since both q and p are treated as coordinates of S, it would only be natural then to minimize the action with respect to both these coordinates to obtain the resultant equations of motion. So we consider the variation $p \to p + \delta p$, which gives the variation in the action (and setting it to zero)

$$\frac{\delta S}{\delta p_j(t)} = \dot{q}_j - \frac{\partial \mathcal{H}}{\partial p_j} = 0. \tag{5.45}$$

A variation in q then gives

$$\frac{\partial S}{\partial q_j(t)} = -\dot{p}_j - \frac{\partial \mathcal{H}}{\partial q_j} = 0, \tag{5.46}$$

both of which once again constitute the Hamilton equations of motion. Not too interesting, but extending this will allow us to rewrite in the action integral in the suggestive form

$$S = \int d\tau \mathcal{L} = \int \left[\sum_j p_j dq_j - \mathcal{H}(p,q) d\tau \right], \tag{5.47}$$

from which Eqs. (5.36, 5.42) are more obviously derived. In a system where the Hamiltonian is conserved, we can utilize Eqs. (5.36, 5.42) to construct a relation from the definition of the Hamiltonian

$$\boxed{-\frac{\partial S}{\partial t} = \left[\sum_j \frac{1}{2m_j} \left(\frac{\partial S}{\partial q_j} \right)^2 + U(q_j) \right]}. \tag{5.48}$$

This is known as the *Hamilton-Jacobi* equation.

Note: The Hamilton-Jacobi equation is the classical mechanics analog of the Schrödinger's equation in quantum mechanics. This can be seen by taking $\Psi = e^{\frac{i}{\hbar}S}$, plugging this solution into the Schrödinger's equation then taking the limit where $\hbar \to 0$.

As an aside, we first note that:

$$F(\boldsymbol{x}) = \int_{\boldsymbol{x}_0}^{\boldsymbol{x}} d\boldsymbol{x}'_j A_j(\boldsymbol{x}') \tag{5.49}$$

is path independent (i.e. $\frac{\partial F}{\partial x_j} = A_j$) **iff**:

$$\frac{\partial A_j}{\partial x_k} = \frac{\partial A_k}{\partial x_j}. \tag{5.50}$$

With this in mind, we can consider again the least action principle and ask if there would be a quantity such that given only the shape of the trajectory and not its parameterization in time, we can equivalently minimize the action. It turns out that this quantity is

$$S_0 = \int \sum_j p_j \dot{q}_j dt = \int \sum_j p_j dq_j, \tag{5.51}$$

often referred to as the *shortened action*, because the variation in S_0 works out to be extremal given that Hamilton's equations are satisfied in a system where the Hamiltonian is conserved. Hamiltonian conservation implies that

$$E = \sum_j \frac{1}{2} m \dot{q}_j^2 + U(q) = \text{constant}, \tag{5.52}$$

$$\Rightarrow \quad S_0 = \int \sqrt{2\left[E - U(\boldsymbol{q})\right] \sum_j m_j (dq_j)^2}. \tag{5.53}$$

In the case where $m_i = m_j = m$ for all $i \neq j$, then the integral simplifies to

$$\boxed{S_0 = \int dL \sqrt{2\left[E - U(\boldsymbol{q})\right]}}, \tag{5.54}$$

where dL is the infinitesimal line segment in multidimensional coordinate space.

§5.5 Canonical Transformations and Poisson Brackets

Going back to the Hamilton equations, we had

$$\dot{q}_j = \frac{\partial \mathcal{H}}{\partial p_j}, \tag{5.55a}$$

$$\dot{p}_k = -\frac{\partial \mathcal{H}}{\partial q_k}. \tag{5.55b}$$

If we abstract away and just look at the mathematical formulation of these relations, we see that it is hard to set q and p apart as variables. That is, the equations which govern the conjugate variables q and p differ simply by a minus sign, which can in fact be absorbed into the Hamiltonian function. Following this vein, we can introduce new functions of these variables $Q_k(q_j, p_j), P_k(q_j, p_j)$, such that we can write a new Hamiltonian $\mathcal{K}(Q, P)$, which gives us the corresponding Hamilton equations

$$\dot{Q}_k = \frac{\partial \mathcal{K}}{\partial P_k}, \tag{5.56a}$$

$$\dot{P}_k = -\frac{\partial \mathcal{K}}{\partial Q_k}. \tag{5.56b}$$

There would still be some freedom in the way these functions are defined, so we can attempt to find functions such that

$$S = \int \left[\sum_k P_k dQ_k - \mathcal{K}(Q_k, P_k) dt \right] = \int \left[\sum_j p_j dq_j - \mathcal{H}(p, q) dt \right] + \int dt \frac{dF}{dt}, \tag{5.57}$$

where F as an arbitrary well-behaved function. Comparing the integrands then give us

$$dF = \sum_j (p_j dq_j - P_k dQ_k) - (\mathcal{K} - \mathcal{H})dt, \tag{5.58}$$

from which we can derive the relations

$$\boxed{\begin{aligned} p_j &= \frac{\partial F}{\partial q_j}, \\ P_j &= -\frac{\partial F}{\partial Q_j}, \\ \mathcal{K} - \mathcal{H} &= \frac{\partial F}{\partial t}. \end{aligned}} \tag{5.59}$$

These are known as the *canonical transformations* and F is known as the *generating function* of the transformation. These relations occur in the particular case where $F = F(\boldsymbol{q}, \boldsymbol{Q}, t)$.

This shows that choosing some function of the form $F(\boldsymbol{q}, \boldsymbol{Q}, t)$ allows use to "generate" relations between the canonical variables, which would then allow us to construct our Hamiltonian. The generalization of these coordinate transformations is 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*:

$$F_1 = F(\boldsymbol{q}, \boldsymbol{Q}, t) \tag{5.60a}$$

$$\Rightarrow \; P_j = \frac{\partial F_1}{\partial Q_j}, \quad p_j = \frac{\partial F_1}{\partial q_j}, \quad \mathcal{K} = \mathcal{H} + \frac{\partial F_1}{\partial t}; \tag{5.60b}$$

2. *Generator Functions of the* 2^{nd} *Kind*:

$$F_2 = F(\boldsymbol{q}, \boldsymbol{P}, t) - \boldsymbol{Q} \cdot \boldsymbol{P} \tag{5.61a}$$

$$\Rightarrow \; Q_j = \frac{\partial F_2}{\partial P_j}, \quad p_j = \frac{\partial F_2}{\partial q_j}, \quad \mathcal{H} = \mathcal{K} + \frac{\partial F_2}{\partial t}; \tag{5.61b}$$

3. *Generator Functions of the* 3^{rd} *Kind*:

$$F_3 = F(\boldsymbol{P}, \boldsymbol{Q}, t) + \boldsymbol{q} \cdot \boldsymbol{p} \tag{5.62a}$$

$$\Rightarrow \; q_j = -\frac{\partial F_3}{\partial P_j}, \quad P_j = -\frac{\partial F_3}{\partial Q_j}, \quad \mathcal{H} = \mathcal{K} + \frac{\partial F_3}{\partial t}; \tag{5.62b}$$

4. *Generator Functions of the 4^{th} Kind*:

$$F_4 = F(\boldsymbol{p}, \boldsymbol{P}, t) - \boldsymbol{Q}\cdot\boldsymbol{P} + \boldsymbol{q}\cdot\boldsymbol{p} \tag{5.63a}$$

$$\Rightarrow\ q_j = -\frac{\partial F_4}{\partial p_j},\quad Q_j = \frac{\partial F_4}{\partial P_j},\quad \mathcal{H} = \mathcal{K} + \frac{\partial F_4}{\partial t}. \tag{5.63b}$$

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.

§5.5.1 Poisson Brackets

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:

$$\begin{aligned}\frac{d}{dt}f(p,q,t) &= \frac{\partial f}{\partial t} + \dot{q}\frac{\partial f}{\partial q} + \dot{p}\frac{\partial f}{\partial p}\\ &= \frac{\partial f}{\partial t} + \frac{\partial \mathcal{H}}{\partial p}\frac{\partial f}{\partial q} - \frac{\partial \mathcal{H}}{\partial q}\frac{\partial f}{\partial p},\\ \Rightarrow\ \frac{d}{dt}f - \frac{\partial}{\partial t}f &= \frac{\partial \mathcal{H}}{\partial p}\frac{\partial f}{\partial q} - \frac{\partial \mathcal{H}}{\partial q}\frac{\partial f}{\partial p} \equiv [\mathcal{H}, f],\end{aligned} \tag{5.64}$$

where $[\mathcal{H}, f]$ is known as the Poisson bracket of the Hamiltonian $\mathcal{H}$ with f (some texts use the notation $\{\ldots,\ldots\}$ instead of $[\ldots,\ldots]$). From this, we see that if the function f we are working with is **not** an explicit function of time, (5.64) reduces to

$$\boxed{\frac{d}{dt}f(p,q) = [\mathcal{H}, f]}, \tag{5.65}$$

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 $f = f(t)$.

$$\frac{d}{dt}f(t) = \frac{\partial}{\partial t}f(t), \tag{5.66}$$

$$\Rightarrow\ [\mathcal{H}, f] = \frac{d}{dt}f(t) - \frac{\partial}{\partial t}f(t) = 0. \tag{5.67}$$

But in this scenario, this does **not** imply that f is conserved. Equation (5.64) actually gave a specific example of a Poisson bracket. We can generalize the definition as follows.

Definition 5.5.1. Poisson Bracket: *Given two functions of conjugate variables $f(\boldsymbol{p}, \boldsymbol{q}, t)$ and $g(\boldsymbol{p}, \boldsymbol{q}, t)$, their Poisson bracket is defined as*

$$[f, g]_{(p,q)} \equiv \sum_j \left(\frac{\partial f}{\partial p_j}\frac{\partial g}{\partial q_j} - \frac{\partial f}{\partial q_j}\frac{\partial g}{\partial p_j}\right) \tag{5.68}$$

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. *Bilinearity*: $[f + g, h] = [f, h] + [g, h]$
3. *Product Rule*: $[f \cdot g, h] = f[g, h] + g[f, h]$
   $\Rightarrow \left[f^2, g\right] = 2f[f, g]$
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. Another very interesting property of Poisson brackets is that they remain invariant under canonical transformations of the conjugate variables. That is to say,

$$[P_j, Q_k]_{(p,q)} = \delta_{j,k}. \tag{5.69}$$

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

$$\frac{df}{dt} = \frac{\partial f}{\partial t} = 0, \quad \frac{dg}{dt} = \frac{\partial g}{\partial t} = 0, \tag{5.70}$$

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

$$\Rightarrow \frac{d}{dt}[f, g] = 0. \tag{5.71}$$

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

$$[\mathcal{H}, f] = [\mathcal{H}, g] = 0. \tag{5.72}$$

Then utilizing the Jacobi identity, we get

$$\begin{aligned} & [\mathcal{H},[f,g]]+[f,[g,\mathcal{H}]]+[g,[\mathcal{H},f]]=0, \\ \Rightarrow\quad & [\mathcal{H},[f,g]]=0, \\ \Rightarrow\quad & \frac{d}{dt}[f,g]=0, \end{aligned} \tag{5.73}$$

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

§5.5.2 Infinitesimal Canonical Transformations

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 2nd 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.

$$F_2' = \boldsymbol{q}\cdot\boldsymbol{P}+\epsilon G(\boldsymbol{q},\boldsymbol{P}). \tag{5.74}$$

Its associated point transformations are then

$$p_j = \frac{\partial F_2'}{\partial q_j} = P_j+\epsilon\frac{\partial G}{\partial q}, \tag{5.75a}$$

$$Q_j = \frac{\partial F_2'}{\partial P_j} = q_j+\epsilon\frac{\partial G}{\partial P_j}. \tag{5.75b}$$

Notice that in the limit as $\epsilon\to 0$, this causes $P_j\to p_j$ and hence,

$$\lim_{\epsilon\to 0}\frac{\partial G(\boldsymbol{q},\boldsymbol{P})}{\partial P_j} = \frac{\partial G(\boldsymbol{q},\boldsymbol{p})}{\partial p_j}. \tag{5.76}$$

Now let $\epsilon = dt$ and $G(\boldsymbol{q},\boldsymbol{P}) = \mathcal{H}(\boldsymbol{q},\boldsymbol{p})$ and substitute these into the point transformation above. As a result, we get

$$\boxed{P_j = p_j - dt\frac{\partial\mathcal{H}}{\partial q_j},\quad Q_j = q_j+dt\frac{\partial\mathcal{H}}{\partial P_j}}, \tag{5.77}$$

$$\begin{aligned} \Rightarrow\quad & P_j = p_j+\dot{p}_j dt,\quad Q_j = q_j+\dot{q}_j dt, \\ \Rightarrow\quad & P_j(t) = p_j(t+dt),\quad Q_j(t) = q_j(t+dt). \end{aligned} \tag{5.78}$$

Hence, we see from (5.77) and (5.78) that the Hamiltonian is indeed the *generator of time translation*. Another way to write this results can be done so in terms of Poisson brackets

$$\boxed{\delta q_j = [\mathcal{H},q_j]\delta t,\quad \delta p_j = [\mathcal{H},p_j]\delta t}, \tag{5.79}$$

where $\delta q_j = q_j - Q_j$ and $\delta p_j = P_j - p_j$. Generalizing this, we can see that if we have a conserved quantity g such that

$$\frac{dg}{dt} = [\mathcal{H}, g] = 0, \tag{5.80}$$

then we can take g to be a generator of a transformation on $\mathcal{H}$ which implies that g **generates a symmetry in the system by means of Poisson brackets**.

§5.6 Solving the Hamilton-Jacobi Equation

In the deep dive section on actions, we derived the Hamilton-Jacobi equation in Eq. (5.48). We will now look to solving this equation (first in 1-dimension). First we consider a canonical transformation with some generator F, such that the new Hamiltonian $\mathcal{K} = 0$. That is,

$$\dot{Q} = \frac{\partial \mathcal{K}}{\partial P} = 0, \tag{5.81a}$$

$$\dot{P} = -\frac{\partial \mathcal{K}}{\partial Q} = 0, \tag{5.81b}$$

$$\mathcal{K} = \mathcal{H} + \frac{\partial F}{\partial t} = 0. \tag{5.81c}$$

This implies that the new conjugate variables P and Q are constants and can be achieved by setting $F = S$, the action. The Hamiltonian $\mathcal{H}$ (without explicit time-dependence), is given by

$$\mathcal{H} = \frac{p^2}{2m} + U(q), \tag{5.82}$$

which gives the Hamilton-Jacobi equation

$$-\frac{\partial S}{\partial t} = \frac{1}{2m}\left(\frac{\partial S}{\partial q}\right)^2 + U(q). \tag{5.83}$$

Using the analogy of quantum mechanics where we consider that the wave function takes the form

$$\Psi = e^{\frac{i}{\hbar}S}, \tag{5.84}$$

and considering an energy eigenstate such that

$$\Psi = \psi e^{-\frac{i}{\hbar}Et}, \tag{5.85}$$

we get that we can consider the action as taking the form

$$S = W - Et, \tag{5.86}$$

where W is a function with no explicit time-independence and E is a constant. Plugging this into the Hamilton-Jacobi equation we get

$$\frac{1}{2m}\left(\frac{\partial W}{\partial q}\right)^2 + U(q) = E, \tag{5.87}$$

which is reminiscent of the time-independent Schrödinger's equation. Rewriting this as an integral equation gives

$$W(E,q) = \int_{q_0}^{q} dq' \sqrt{2m[E - U(q')]}. \tag{5.88}$$

Treating E as the new coordinate after canonical transformation (that is $Q = E$, by assertion that we are looking for a Q that is a constant of motion), we have that the new momentum is then

$$\boxed{P(E,q) = \frac{\partial S}{\partial E} = \left[\int_{q_0}^{q} \frac{dq'}{\sqrt{[E - U(q')]}}\right] - t}. \tag{5.89}$$

From this, we can then solve for $q(t)$ in terms of P and E which are constants (i.e. set by initial conditions). As for higher dimensional systems, it is in general very complicated to generalize the formalism we have above. However, if the Hamiltonian works out to be separable such that

$$\mathcal{H} = \mathcal{H}[q_1, p_1, h(q_j, p_j)]; \tag{5.90a}$$

$$\text{where} \quad h(q_j, p_j) = \alpha, \quad \mathcal{H} = E, \tag{5.90b}$$

with α and E being constants, then the action can also be separated into the terms

$$S = S_1(q_1, \alpha, E) + \tilde{S}(q_j, \alpha) - Et. \tag{5.91}$$

So we see that the motion in this system is now distinctly reducible into 2 separate sub-systems which grant the equations of motion generated by

$$P_1 = \frac{\partial S}{\partial \alpha}, \tag{5.92a}$$

$$P = \frac{\partial S}{\partial E}, \tag{5.92b}$$

which allow us to solve for $q_1(t)$ and a family of solutions to $q_j(t)$ which lie on the level set $h(q_j, p_j) = \alpha$.

Note: Systems in which the variables are separable up to the number of degrees of freedom are known as *integrable systems*. The name stems from the fact that α and E (or in general a set of $\{\alpha_j\}$) are integrals of motion which are in functions of the conjugate variables that remain constant throughout dynamics evolution.

§5.6.1 Action-Angle Variables

For a 1-dimensions system, having the transformed coordinate $Q = E$ was a convenient choice which led to a solution in quadratures. However, extending this to a general multidimensional system naively could result in a number of complications. A better set of variables are ones which encode the integrals of motion into them. Specifically, if we are dealing with systems that

have closed trajectories (a.k.a closed orbits) or even just orbits that are bounded (confined to a particular region but not necessarily closed), we can define the *action variables* I, defined as the area enclosed by this closed trajectory in phase space:

$$\boxed{I = \frac{1}{2\pi}\oint p dq}. \tag{5.93}$$

This variable is a constant of motion and can be mapped to the energy of that closed trajectory, which implies $\mathcal{H} = \mathcal{H}(I)$. We take the action variable as the canonical transform momentum which then leaves the canonical transform position w, to be defined. Recall the shortened action S_0, defined in Eq. (5.51). We use this to define w as

$$w \equiv \frac{\partial S_0}{\partial I}, \tag{5.94}$$

which implies that

$$\dot{I} = -\frac{\partial \mathcal{H}}{\partial w} = 0, \tag{5.95a}$$

$$\dot{w} = \frac{\partial \mathcal{H}}{\partial I} \equiv \omega(I), \tag{5.95b}$$

where $\omega(I)$ is a constant. So we have that

$$\boxed{w(t) = \omega t + \varphi = \left(\frac{\partial \mathcal{H}}{\partial I}\right) t + \varphi}, \tag{5.96}$$

which looks very much like an angle, hence we call it the *angle variable*. In fact, w is some function of the angle which sweeps the closed trajectory such that it varies linearly with time and changes by 2π over one orbit. So we can say that I tells us which allowable closed trajectory we picked, while w tells us where on that trajectory we are. The action-angle variable formalism allows us to think of all periodic trajectories as some form of harmonic oscillator type motion, which is a very intuitive picture of things.

In a multidimensional (N degree of freedom) system with such closed orbits, we know that Hamiltonian conservation would already restrict the motion in phase space to the $2N-1$ dimensional energy manifold defined by $\mathcal{H}(q_j, p_j) = E$. The additional assertion of closed orbits then further restricts the motion to generalized torus-like manifolds (a.k.a invariant tori). Even within such surfaces, the orbits can in general be characteristically quite different. Specifically, given a system with bounded orbits, we can have that the orbits are either:

1. bounded and closed: if we have action-angle variables such that $\omega_i/\omega_j \in \mathbb{Q}$ (rational), implying that we can find $\omega_j t = 2\pi n_j$ with $n_j \in \mathbb{Z}$;
2. bounded but not closed: if we have action-angle variables such that $\omega_i/\omega_j \notin \mathbb{Q}$ (irrational).

In fact, bounded and closed orbits are rather special cases as for instance discussed in Thm. 1.4.1.

Note: There is a theorem known as the *Kolmogorov–Arnold–Moser theorem* which states that even under sufficiently small but possibly nonlinear perturbations, invariant tori can

remain tori given that it satisfies the *non-resonance condition* (no non-singular perturbation terms). Otherwise, the tori will be *destroyed* and the system transits to one that is non-integrable.

If a system as exactly N independent integrals of motion defined by

$$g_n(q,p) = \text{constant}, \tag{5.97a}$$
$$\text{s.t.} \quad \{\mathcal{H}, g_n\} = 0, \tag{5.97b}$$

then the system is said to be *integrable* which nicely solves the Hamilton-Jacobi equation. These integrals of motion define the manifold in phase space which the motion of the system is confined to. If however, we were to add nonlinear interaction terms to such systems, this could in principle lead to chaotic behavior which we will now study.

§5.7 Classical Chaos

Chaotic systems lie in the interim between well-behaved integrable systems and completely random ones, for which chaotic dynamics arises from nonlinearities in a dynamical system. This is often seen to emerge when some parameters of the system are tuned past certain numerical values. Because of its non-analytic nature, it is often difficult to generalize studies of chaotic system (at least without having to delve into the deep mathematical formalism of symplectic manifolds) beyond the use of examples to illustrate key concepts. Arguably the simplest such example we can start with is of a damped-driven pendulum with equation of motion

$$\ddot{\phi} + 2\beta\dot{\phi} + \omega_0^2 \sin\phi = \gamma\omega_0^2 \cos(\omega t) \tag{5.98}$$

Choosing the initial conditions and parameters

$$\phi(0) = \frac{\pi}{4}, \quad \dot{\phi}(0) = 0, \tag{5.99a}$$
$$\omega_0 = 3\pi, \quad \omega = 2\pi, \quad \beta = \frac{\omega_0}{2}, \quad \gamma = 0, \tag{5.99b}$$

we see that the system just spirals inward in the phase plane as shown in Fig. 5.3. The blue point in the figure is the end point of the trajectory. Now when we turn on the driving ($\gamma = 0.2$), we see that it transits to motion which looks like that in Fig. 5.4. Interestingly, we still see that the trajectory continues to spiral toward the $(\phi, \dot{\phi}) = (0, 0)$ point in the phase plane. In fact, this occurs even for different initial value chosen for $\phi(0)$ and $\dot{\phi}(0)$! We call such a point an *attractor*, in which for a wide range of initial configurations, our system will always evolve in phase space towards the attractor.

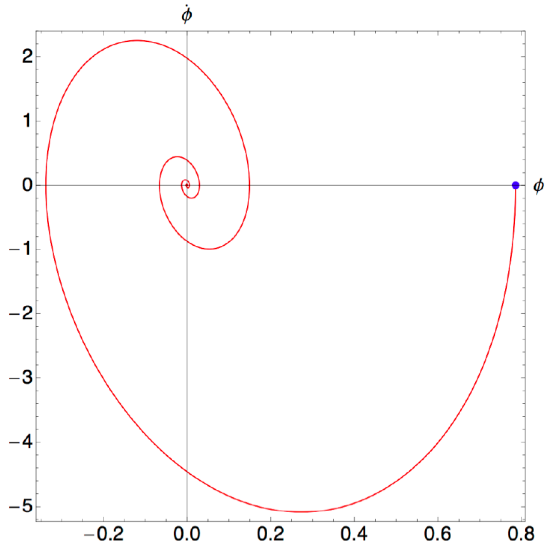

Figure 5.3: Phase plane trajectory of an undriven damped oscillator.

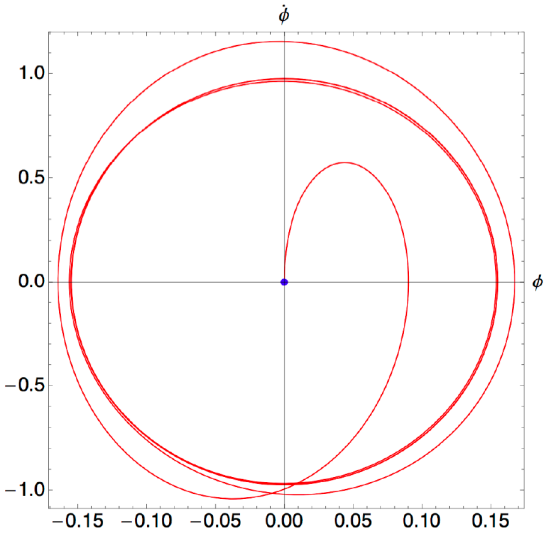

Figure 5.4: Phase plane trajectory of a driven damped oscillator ($\gamma = 0.2$).

§5.7.1 Period Doubling

If we now vary the strength of the driving, we find that the system adopts a new limiting behavior after sufficient evolution in time around $\gamma = 1.07$. That is, the motion now tends toward a periodic orbit known as a *limit cycle* as shown in Fig. 5.5

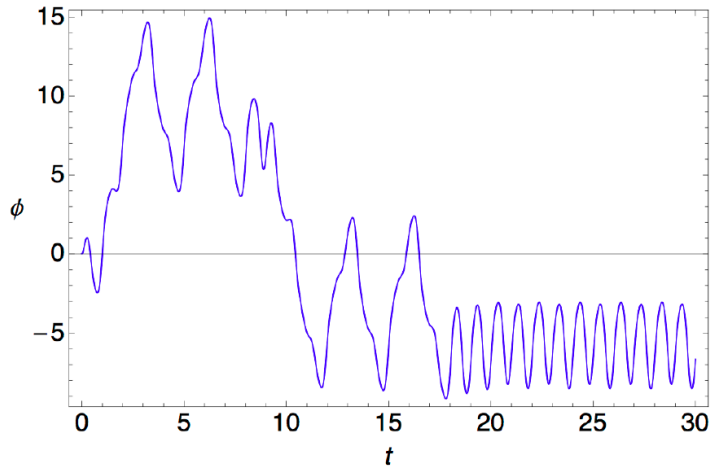

Figure 5.5: $\phi(t)$ vs t trajectory of a driven damped oscillator ($\gamma \approx 1.07$).

Even more intriguing is that if we look closely at this plot this on the phase plane (Fig. 5.6), we notice 2 distinct cycles emerge that alternate between one another.

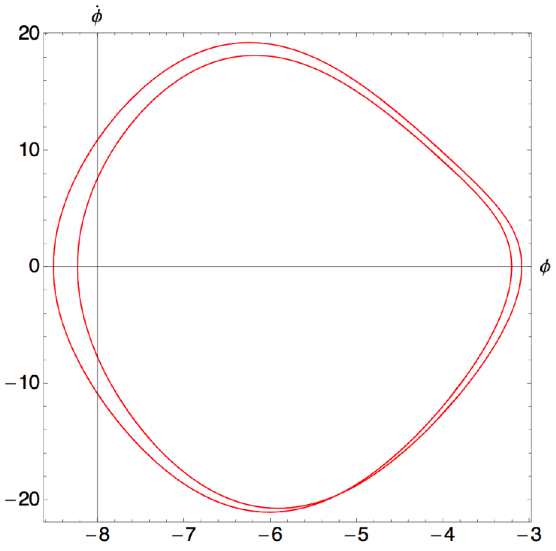

Figure 5.6: Phase plane trajectory of a driven damped oscillator ($\gamma \approx 1.07$).

If we look at the period of oscillations, system is no longer oscillating with period 1 (set by the drive frequency), but now has twice the expected period. This phenomenon is known as *period doubling*. This is not chaotic motion but this is something which occurs en route to chaos. This period doubling phenomena in fact occurs for several values of γ:

γ	Period
1.0663	$1 \to 2$
1.0793	$2 \to 4$
1.0821	$4 \to 8$
1.0827	$8 \to 16$
$\vdots$	$\vdots$

Table 5.1: Period doubling occurrences.

which in fact follows the relation

$$\gamma_{n+1} - \gamma_n = \frac{1}{\delta}(\gamma_n - \gamma_{n-1}), \tag{5.100}$$

with $\delta \approx 4.669$ (a transcendental number like e or π) known as the *Feigenbaum number*. The Feigenbaum number is in fact universal and occurs in any period doubling nonlinear system. Another very illuminating picture that can be plotted is the so-called *bifurcation diagram*, in which the steady-state value of $\phi(t)$ is taken along the $\dot{\phi} = 0$ plane, and plot as a function of γ (shown in Fig. 5.7).

Note: The plot of points representing all intersection of the trajectory through a planar section is known as a *Poincaré recurrence map*. So a single slice of the bifurcation diagram for a particular value of γ is in fact a Poincaré recurrence map.

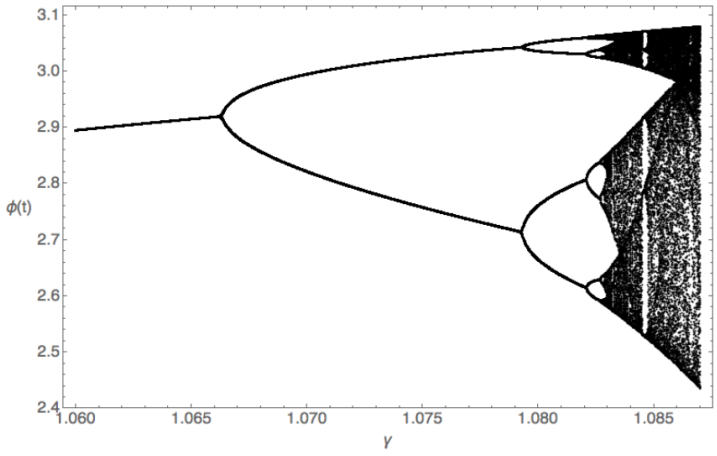

Figure 5.7: Bifurcation diagram of a driven damped oscillator as a function of γ.

This diagram beautiful illustrates the "route to chaos" of the system as γ is tuned, for which we see that period doubling occurs as chaos (completely aperiodic evolution) is approached. A quantitative measure use to determine if a trajectory is chaotic is known as the *Lyapunov exponent*. It was observed by Aleksandr Lyapunov, that the absolute difference in trajectories $|\Delta\phi(t)|$ when evolved from different initial conditions followed exponential trends. That is, for

systems with stable attractors and limit cycles, the quantity $|\Delta\phi|$ would exponentially decay, whereas for chaotic systems, $|\Delta\phi|$ would exponentially increase:

$$\text{non-chaotic (stable):} \quad |\Delta\phi(t)| \sim e^{-\lambda t}, \tag{5.101a}$$

$$\text{chaotic:} \quad |\Delta\phi(t)| \sim e^{+\lambda t}, \tag{5.101b}$$

where λ is a positive constant referred to as the Lyapunov exponent. The expression above are mediated by $\sim$ instead of an equality because there might be some other multiplicative function (possibly stochastic) attached to the exponential, which generally we are not concerned with. It is only the exponential trend we look out for to characterize dynamical systems as chaotic.

§5.7.2 Iterative Maps

Classical chaos is often seen to arise not just in continous dynamical systems, but also discrete ones. That is, system of the form

$$\boldsymbol{q}_{n+1} = \boldsymbol{f}(\boldsymbol{q}_n, \lambda), \tag{5.102}$$

where $\boldsymbol{f}$ is some nonlinear function of the previous state and λ is the chaos parameter(s). These functions are what are known as *iterative maps*, that take a current state to the next state stepwise. This in practice, is in fact often a more accurate representation of a dynamical system because sampling of the system state is usually performed stroboscopically. This results in a *difference equation* as opposed to a differential equation. A famous example of an iterative map, is known as the *logistic map* written as

$$f(x_n, r) = r x_n (1 - x_n). \tag{5.103}$$

In the expression above, we replace $\boldsymbol{q} \to x$ and $\lambda \to r$. This results in a bifurcation diagram as shown in Fig. 5.8.

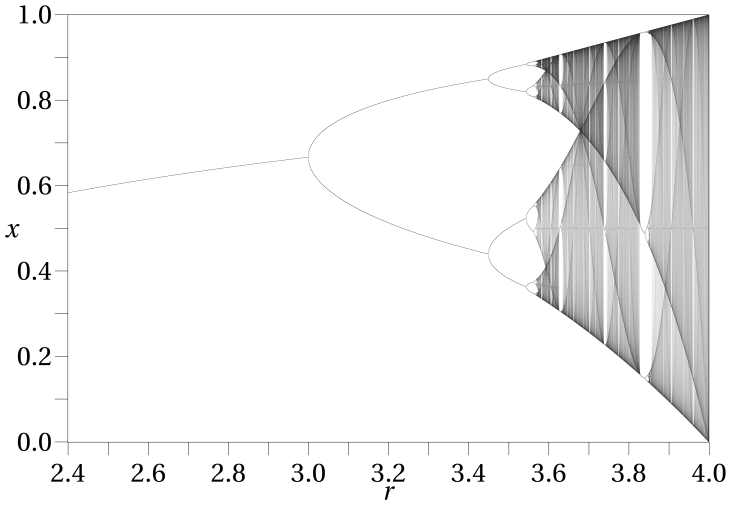

Figure 5.8: Bifurcation diagram of a logistic map.

With iterative maps, attractors can be found via the condition that

$$\left|\frac{\partial q_{n+1}}{\partial q_n}\right|_{q_n \approx \bar{q}} < 1, \tag{5.104}$$

where $\bar{q}$ is the supposed value of the attractor. The statement is telling us that in the vicinity of the attractor, the system should evolve in such a way that its trajectory tends/remains bounded to this attractor. This absolute valued measure is of course a function of the chaos parameter λ, which we will see change from < 1 to > 1 when the parameter is varied into the chaotic regime.

§5.8 Classical Field Theory

What is a field? So far, we have studied the dynamics of particles which have been specified by some set of generalized coordinates in phase space. Particles are thus discrete and localized objects that traverse through space. On the other hand, fields are extended objects which exist at every point in space at any instant in time. Because of this, it is not possible to characterize a field by its position through some generalized coordinate, but rather we will need to specify the "strength" of the field at every point in space. The corresponding analog of the velocity of fields would then be the rate of change of the field strength at every point in space. Historically, the concept of a field was first put forth by Faraday pertaining to the electromagnetic field. However, we shall not be using this as the starting point but instead, utilize the continuum (scalar) field description of elastic mediums.

For sufficiently large objects, we can treat a solid with finite elastic modulus as a continuous medium (i.e. field). To see this, we first consider a system of N coupled spring-masses along 1-dimension separated at equilibrium by distance a. Each of these masses will be at position

$$x_n = na + u_n, \tag{5.105}$$

where u_n is taken as an out-of-equilibrium displacement. The Lagrangian is thus

$$\mathcal{L} = \sum_n \frac{1}{2} m\dot{u}_n^2 - \sum_n \frac{1}{2} k(u_{n+1} - u_n)^2, \tag{5.106}$$

which grant the equations of motion

$$m\ddot{u}_n = k(u_{n+1} + u_{n-1} - 2u_n), \tag{5.107}$$

via the Euler-Lagrange equations. In the limit where $N \to \infty$ and $a \to 0$, we can now introduce the concept of a field defined as

$$\phi(na) = u_n, \tag{5.108}$$

which in words means that the value of the field at position na is equal to the displacement of

the mass originally at position na. So the Lagrangian becomes

$$\begin{aligned}
\lim_{a\to 0}\mathcal{L} &= \lim_{a\to 0}\left[\frac{m}{2}\sum_n \dot{\phi}^2 - \frac{k}{2}\sum_n \left[\phi((n+1)a) - \phi(na)\right]^2\right] \\
&\approx \lim_{a\to 0}\left[\frac{m}{2}\sum_n \dot{\phi}^2 - \frac{k}{2}\sum_n \left(a\left.\frac{\partial\phi}{\partial x}\right|_{x=na}\right)^2\right] \\
&\approx \lim_{a\to 0}\int dn\left[\frac{m}{2}\dot{\phi}^2 - \frac{k}{2}\left(a\left.\frac{\partial\phi}{\partial x}\right|_{x=na}\right)^2\right] \\
&= \lim_{a\to 0}\int \frac{dx}{a}\left[\frac{m}{2}\dot{\phi}^2 - \frac{k}{2}\left(a\frac{\partial\phi}{\partial x}\right)^2\right] \\
&= \int dx\left[\frac{\rho}{2}\dot{\phi}^2 - \frac{\sigma}{2}\left(\frac{\partial\phi}{\partial x}\right)^2\right],
\end{aligned} \tag{5.109}$$

where $\rho = \lim_{a\to 0} m/a$ and $\sigma = \lim_{a\to 0} ka$. We call the integrand the *Lagrangian density* $\mathcal{L}$, and we now call the total integral the *Lagrange function* L. So the action is then

$$S = \int dt L = \int dt \int d^3x \mathcal{L}. \tag{5.110}$$

Note: Systems in which a Lagrangian density $\mathcal{L}$, can be explicitly written down are referred to as *local field theories*. In general, there are theories in which only the Lagrangian L can be written down, which would then constitute a *non-local field theory*.

For fields, the resulting Euler-Lagrange equations can be derived once again via the calculus of variations to give

$$\boxed{\frac{d}{dt}\frac{\partial\mathcal{L}}{\partial\dot{\phi}} + \frac{d}{d\boldsymbol{x}}\cdot\frac{\partial\mathcal{L}}{\partial(\partial\phi/\partial\boldsymbol{x})} = \frac{\partial\mathcal{L}}{\partial\phi}}. \tag{5.111}$$

In the more compact index notation, this is simply written as

$$\boxed{\partial_\mu\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)} = \frac{\partial\mathcal{L}}{\partial\phi}}, \tag{5.112}$$

where $\mu = t, x, y, z$ and repeated indices are summed over (these are Euclidean sums). Going back to our example of the elastic medium, we had the Lagrangian density as

$$\mathcal{L} = \frac{1}{2}\rho\dot{\phi}^2 - \frac{1}{2}\sigma\left(\frac{\partial\phi}{\partial x}\right)^2, \tag{5.113}$$

which when plugged back into the Euler-Lagrange equations gives us

$$\rho\frac{\partial^2\phi}{\partial t^2} = \sigma\frac{\partial^2\phi}{\partial x^2}, \tag{5.114}$$

which is in fact the wave equation! In 3-spatial dimensions, we instead get

$$\rho \frac{\partial^2 \phi}{\partial t^2} = \sigma \boldsymbol{\nabla}^2 \phi. \tag{5.115}$$

§5.8.1 Generalized Field Variables and Poisson Brackets

As in classical analytical mechanics, we also want to obtain a set of generalized variables that describe the system. We have already seen that the analog of generalized coordinates in field theories are the fields themselves, but what about the momenta? For field theories, we will call the momentum π, and define it as

$$\boxed{\pi(\boldsymbol{r}) \equiv \frac{\delta \mathcal{L}}{\delta \dot{\phi}(\boldsymbol{r})}}, \tag{5.116}$$

with δ denoting the variational derivative. This definition although accurate, might be a little cumbersome and so for **local** fields, we can define the momentum instead as

$$\pi(\boldsymbol{r}) = \frac{\partial \mathcal{L}}{\partial \dot{\phi}(\boldsymbol{r})}. \tag{5.117}$$

The Hamiltonian for field theories is then defined as

$$\boxed{H \equiv \int d^3x \left[\pi(\boldsymbol{r}) \dot{\phi}(\boldsymbol{r}) - \mathcal{L} \right]}, \tag{5.118}$$

where we can also call the integrand as the *Hamiltonian density*, $\mathcal{H}$. In terms of the Hamiltonian, we can also have the analogous Hamilton equations for fields as

$$\dot{\pi} = -\frac{\delta H}{\delta \phi(\boldsymbol{r})}, \tag{5.119a}$$

$$\dot{\phi} = \frac{\delta H}{\delta \pi(\boldsymbol{r})}, \tag{5.119b}$$

Example:

Going back to our initial example on the continuous extrapolation of coupled oscillators, we can obtain that the Hamiltonian density for that system is given by

$$\mathcal{H} = \frac{\pi^2}{2\rho} + \frac{\sigma}{2} \left(\frac{\partial \phi}{\partial x} \right)^2, \tag{5.120}$$

which then results in the Hamilton equations:

$$\dot{\phi} = \frac{\partial \mathcal{H}}{\partial \pi} = \frac{\pi}{\rho}, \tag{5.121a}$$

$$\dot{\pi} = \rho \ddot{\phi} = -\frac{\delta H}{\delta \phi} = \sigma \frac{\partial^2 \phi}{\partial x^2}, \tag{5.121b}$$

where the second of the 2 equations above is once again the wave equation.

Additionally, we can also define poisson brackets for field variables as:

$$\boxed{\{A, B\} = \int d^3x \left[\frac{\delta A}{\delta \pi(\boldsymbol{r})} \frac{\delta B}{\delta \phi(\boldsymbol{r})} - \frac{\delta B}{\delta \pi(\boldsymbol{r})} \frac{\delta A}{\delta \phi(\boldsymbol{r})} \right]}. \tag{5.122}$$

It is also useful to note the relation

$$\frac{\delta \phi(\boldsymbol{r}_1)}{\delta \phi(\boldsymbol{r}_2)} = \delta(\boldsymbol{r}_1 - \boldsymbol{r}_2). \tag{5.123}$$

§5.8.2 Symmetries and Conservation Laws

As seen before (Sec. 1.2), Noether's theorem also holds field theories in which symmetries generated by infinitesimal space and time translations (along with others) lead to conserved momenta and energy (etc) in the system. However in addition to this, we can also consider an infinitesimal variation of the field

$$\phi \quad \rightarrow \quad \phi + \delta\phi, \tag{5.124}$$

which if the action is invariant to, leads to what are known as *gauge symmetries*. With symmetries, we of course know that there is also an associated Noether current that is conserved. To see this, we first note that the field theory is invariant to variations in the Lagrangian density up to a 4-divergence (divergence of 3-space and 1-time dimension):

$$\mathcal{L}(t, \boldsymbol{x}) \quad \rightarrow \quad \mathcal{L}(t, \boldsymbol{x}) + \partial_\mu \Lambda^\mu(t, \boldsymbol{x}), \tag{5.125}$$

Now considering the variation of the field $\phi \rightarrow \phi + \delta\phi$, we see that this changes the Lagrangian as follows:

$$\begin{aligned} \delta\mathcal{L} &= \frac{\partial \mathcal{L}}{\partial \phi} \delta\phi + \frac{\partial \mathcal{L}}{\partial (\partial_\mu \phi)} \partial_\mu (\delta\phi) \\ &= \left[\frac{\partial \mathcal{L}}{\partial \phi} - \partial_\mu \frac{\partial \mathcal{L}}{\partial (\partial_\mu \phi)} \right] \delta\phi + \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial (\partial_\mu \phi)} \delta\phi \right) \\ &= \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial (\partial_\mu \phi)} \delta\phi \right), \end{aligned} \tag{5.126}$$

where we noticed that the term in the square-bracket vanishes due to the Euler-Lagrange equation. Comparing terms in the variation of the Lagrangian density, we see that

$$\begin{aligned} & \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial (\partial_\mu \phi)} \delta\phi \right) - \partial_\mu \Lambda^\mu(t, \boldsymbol{x}) = 0, \\ \Rightarrow \quad & \partial_\mu \left[\frac{\partial \mathcal{L}}{\partial (\partial_\mu \phi)} \delta\phi - \Lambda^\mu(t, \boldsymbol{x}) \right] = 0, \end{aligned} \tag{5.127}$$

which gives us that the term in square-brackets above is conserved:

$$\boxed{j^\mu = \frac{\partial \mathcal{L}}{\partial (\partial_\mu \phi)} \delta\phi - \Lambda^\mu(t, \boldsymbol{x})}, \tag{5.128}$$

referred to as the *Noether current*, that satisfies the *continuity equation* $\partial_\mu j^\mu = 0$.

Note: The raising and lowering of indices here bear no weight since we are working in Euclidean 4-space. However in relativistic field theories, these play an important role for which we adopt the relativistic conventions here for consistency.

Writing the continuity equation explicitly in space and time components gives

$$\frac{\partial}{\partial t} j^0 + \boldsymbol{\nabla} \cdot \boldsymbol{j} = 0. \tag{5.129}$$

We can integrate this equation over some finite region Ω, to get

$$\begin{aligned} \int_\Omega \frac{\partial}{\partial t} j^0 dV &= -\int_\Omega \boldsymbol{\nabla} \cdot \boldsymbol{j} dV \\ &= -\int_{\partial\Omega} d\boldsymbol{s} \cdot \boldsymbol{j}, \end{aligned} \tag{5.130}$$

where the surface integral (from the divergence theorem) is then the flux through the boundary of region Ω.

§5.8.3 The Stress-Energy Tensor

Going back to the symmetries in space and time coordinates, we can now consider a conserved quantity which arises due to both of these having symmetry transformations. Combining both space and time into a vector indexed by Greek indices x_μ, we consider the transformation

$$x_\mu \quad \to \quad x_\mu + \epsilon_\mu(x). \tag{5.131}$$

Plugging this into the action, we then have

$$\delta S = \int d^4x \mathcal{L}[\phi(x_\mu + \epsilon_\mu), \tilde{\partial}_\mu \phi(x_\mu + \epsilon_\mu)] - \int d^4x \mathcal{L}[\phi(x_\mu), \partial_\mu \phi(x_\mu)]. \tag{5.132}$$

To compute this, we require a change of variable which will result in the Jacobians

$$\begin{aligned} \det\left(\frac{\partial x_\mu}{\partial x_\nu} + \frac{\partial \epsilon_\mu}{\partial x_\nu}\right) &= \det\left(\delta_{\mu\nu} + \frac{\partial \epsilon_\mu}{\partial x_\nu}\right) \\ &= \exp\left[\mathrm{Tr}\ln\left(\delta_{\mu\nu} + \frac{\partial \epsilon_\mu}{\partial x_\nu}\right)\right] \\ &\approx 1 + \frac{\partial \epsilon_\mu}{\partial x_\mu}, \end{aligned} \tag{5.133}$$

and the partial derivative becomes

$$\tilde{\partial}_\mu = \sum_\nu \left(\delta_{\mu\nu} - \frac{\partial \epsilon_\nu(x)}{\partial x_\mu}\right). \tag{5.134}$$

With these, the variation in the action then becomes

$$\begin{aligned}\delta S &= \int d^4x \left[\frac{\partial \epsilon_\mu}{\partial x_\mu}\mathcal{L} - \frac{\partial \mathcal{L}}{\partial(\partial_\mu\phi)}\frac{\partial \epsilon^\nu}{\partial x^\mu}\partial_\nu\phi\right] \\ &= \int d^4x \left[\partial_\mu\left(\delta^{\mu\nu}\epsilon_\nu\mathcal{L} - \epsilon^\nu\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\partial_\nu\phi\right) - \epsilon_\nu\left(\delta^{\mu\nu}\partial_\mu\mathcal{L} - \partial_\mu\left[\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\partial_\nu\phi\right]\right)\right].\end{aligned} \tag{5.135}$$

By the Euler-Lagrange equation and taking that ϵ vanishes on the boundary, the second integrand term above vanishes, leaving the conserved current as

$$j^\mu = \left[\delta^{\mu\nu}\mathcal{L} - \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\delta^{\nu\beta}\partial_\beta\phi\right]\epsilon_\nu. \tag{5.136}$$

To extend this, we can define a tensor

$$\boxed{T^{\mu\nu} \equiv \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\delta^{\nu\beta}\partial_\beta\phi - \delta^{\mu\nu}\mathcal{L}}, \tag{5.137}$$

referred to as the *stress-energy tensor* that is independent of ϵ, that gives rise to 4 conserved currents such that

$$\partial_\mu T^{\mu\nu} = 0. \tag{5.138}$$

The reason for the name of this tensor, is that the T^{00} entry is in fact the energy density, T^{0j} are the momentum densities and all other components are energy fluxes (T^{j0}) and momentum fluxes (T^{jk}). So we have that energy and momentum conservation are given in terms of the stress-energy tensor as

$$\text{energy conservation}: \quad \partial_\mu T^{\mu 0} = 0, \tag{5.139a}$$

$$\text{momentum conservation}: \quad \partial_\mu T^{\mu j} = 0. \tag{5.139b}$$

The spatial components of the stress-energy tensor $T^{jk} = \sigma^{jk}$, are also known as stresses.

Note: Relativistic theories result in stress-energy tensors that are symmetric.

In practice, stress-energy tensors also have an ambiguity in its definition because if we add another tensor term $\partial_\alpha B^{\alpha\mu\nu}$ to $T^{\mu\nu}$ where

$$B^{\alpha\mu\nu} = -B^{\mu\alpha\nu}, \tag{5.140a}$$

$$B^{\alpha\mu\nu} = B^{\alpha\nu\mu}, \tag{5.140b}$$

this still results in all necessary conservation laws associated to $T^{\mu\nu}$, since $\partial_\mu\partial_\alpha B^{\alpha\mu\nu} = 0$.

Appendices

Appendix A

Functionals

Functions are maps which take in functions and return numbers

$$F[f(x)] : f(x) \mapsto a. \tag{A.1}$$

The next step would be to ask, how does one then take derivatives of such object? In order to do so, it would be useful to first refresh our memory of δ-functions.

§A.1 Delta Functions

δ-functions are objects often encountered in quantum mechanics, but has applications to every field of physics. There are 2 types of δ-functions which handle discrete and continuous problems. The discrete version is known as the *Kronecker δ-function*, defined as

$$\delta_{i,j} = \begin{cases} 1, & i = j \\ 0, & \text{otherwise.} \end{cases} \tag{A.2}$$

The continuous version is known as the *Dirac δ-function*, most conveniently thought of in terms of its limit definition of a Lorentzian

$$\delta(x) = \lim_{\varepsilon \to 0} \frac{1}{\pi} \frac{\varepsilon^2}{x^2 + \varepsilon^2}. \tag{A.3}$$

With this, one can easily check that this function satisfies the properties

$$\delta(x) = \begin{cases} \infty, & x = 0 \\ 0, & \text{otherwise,} \end{cases} \tag{A.4a}$$

$$\int_{-\infty}^{\infty} \delta(x) dx = 1. \tag{A.4b}$$

This limit definition allows us to take derivatives of the δ-function as well, since Lorentzians are continuous functions and so we can first take the derivative of the Lorentzian and then

take the limit as $\varepsilon \to 0$. This grants that the resulting derivations of the δ-function follow the relation

$$\int_{-\infty}^{\infty} \frac{d^n \delta(x)}{dx^n} f(x) dx = (-1)^n \frac{d^n f(x)}{dx^n}\bigg|_{x=0}. \tag{A.5}$$

§A.2 Derivatives of Functionals

At this point, we then consider a derivative of a functional with respect to a function argument, written in its limit definition as

$$\frac{\delta F[f(x)]}{\delta f(x^*)} = \lim_{\varepsilon \to 0} \frac{F[f(x) + \varepsilon \delta(x - x^*)] - F[f(x)]}{\varepsilon}, \tag{A.6}$$

with x^* being a particular value of x where we are varying the function at. Setting up a derivative to functionals allows us to then do things like take Taylor expansions of functionals. For instance, consider a "small" perturbative function $\eta(x)$ to $f(x)$. We can then Taylor expand the functional around $f(x)$, written to first-order as

$$F[f(x) + \eta(x)] \approx F[f(x)] + \int dx' \frac{\delta F[f(x)]}{\delta f(x')} \eta(x') + \dots. \tag{A.7}$$

Let's look at an application of functional derivatives, by employing it to derive the Euler-Lagrange equations of motion.

§A.2.1 Euler-Lagrange Equations from Functional Derivatives

Starting with the action (which is a functional) for a one-degree of freedom system,

$$S = \int_{t_i}^{t_f} dt \mathcal{L}(q, \dot{q}), \tag{A.8}$$

we consider the derivative of the action with respect to $q(t)$, written as

$$\begin{aligned}
\frac{\delta S[q(t)]}{\delta q(t')} &= \lim_{\varepsilon \to 0} \frac{\int_{t_i}^{t_f} dt \mathcal{L}\left(q(t) + \varepsilon \delta(t - t'), \dot{q}(t) + \varepsilon \dot{\delta}(t - t')\right) - \int_{t_i}^{t_f} dt \mathcal{L}(q(t), \dot{q}(t))}{\varepsilon} \\
&= \lim_{\varepsilon \to 0} \frac{\int_{t_i}^{t_f} dt \left[\frac{\partial \mathcal{L}}{\partial q(t)} \varepsilon \delta(t - t') + \frac{\partial \mathcal{L}}{\partial \dot{q}(t)} \varepsilon \dot{\delta}(t - t')\right]}{\varepsilon} \\
&= \frac{\partial \mathcal{L}}{\partial q(t')} + \frac{\partial \mathcal{L}}{\partial \dot{q}(t)} \delta(t - t')\bigg|_{t_i}^{t_f} - \int_{t_i}^{t_f} dt \frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \dot{q}(t)} \delta(t - t') \\
&= \frac{\partial \mathcal{L}}{\partial q(t')} - \frac{d}{dt'} \frac{\partial \mathcal{L}}{\partial \dot{q}(t')}.
\end{aligned} \tag{A.9}$$

To extremize S then, we set the first derivative to zero which indeed leaves us with the Euler-Lagrange equation

$$\frac{\partial \mathcal{L}}{\partial q(t)} - \frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \dot{q}(t)} = 0. \tag{A.10}$$

It is then easy to generalize this result to a system with multiple degrees of freedom.