30.104 Dynamics

SUTD, Fall 2018

Dario Poletti Reuben R.W. Wang

Important dates

- Wednesday/Friday of week 3 → 1D project decided
- Wednesday afternoon of week 6 → Midterm test ☺ (2 hours)
- Thursday of week 8 → 1D prototype ready
- Friday of week 8 → make-up class for a PH
- Week 9 → 1D (report due Friday)
- Friday of week 13 → Matlab Exam ☺ (90 mins)
- Friday of week 14 → Final Exam ☺ (2.5 hours)

Figure 1: Important Dates

Evaluation

- Midterm + Final → 55%
- Numerical simulations 5%
- 1D → 25%
- Homework → 10%
- Attendance and Participation → 5%

Figure 2: Course Evaluation

Instructor: *Professor Dario Poletti.*
Instructor office hours: *Wednesday: 1530 - 1630 (3.401.04).*
Instructor email: `dario_poletti@sutd.edu.sg`.

Contents

Chapter 1

Introductory Kinematics

"In this class I hope you will learn not merely results, or formulae applicable to cases that may possibly occur in our practice afterwards, but the principles on which those formulae depend, and without which the formulae are mere mental rubbish. I know the tendency of the human mind is to do anything rather than think. But mental labour is not thought, and those who have with labour acquired the habit of application, often find it much easier to get up a formula than to master a principle"

— J.C. Maxwell, 1860

§1.1 Newtonian Mechanics

We know from basic classical mechanics, Newton's Second Law states that $\vec{F} = \frac{d\vec{p}}{dt}$. Today, we will take a closer look at acceleration.

Definition 1.1.1. Acceleration: *The acceleration of a body is defined as the second derivative of its position with respect to time*

$$\vec{a} = \frac{d^2\vec{r}}{dt^2}. \tag{1.1}$$

Consider the differential motion of a **single** particle from $\vec{A}$ to $\vec{A}'$. Setting a Cartesian coordinate system and having the x and y orthonormal unit vectors as denoted by e_x and e_y ($\vec{e}_i \cdot \vec{e}_j = \delta_{ij}$), we can denote any arbitrary position $\vec{r} = x\vec{e}_x + y\vec{e}_y$. Then the infinitesimal change in position of the particle from $\vec{A}$ to $\vec{A}'$ can be written as $d\vec{r} = dx\ \vec{e}_x + dy\ \vec{e}_y$. It then follows that the velocity and accelerations are respectively:

$$\frac{d\vec{r}}{dt} = \dot{x}\vec{e_x} + \dot{y}\vec{e}_y \tag{1.2}$$

$$\frac{d^2\vec{r}}{dt^2} = \ddot{x}\vec{e_x} + \ddot{y}\vec{e}_y \tag{1.3}$$

§1.2 Time-Varying Coordinates

However, here we have made an implicit assumption! That is, the frame is time-invariant (does not change with time). In fact, it is sometimes even more convenient to make use of time varying coordinate systems. For instance, let's say we are in some rotating frame (rotating in the x, y-plane).

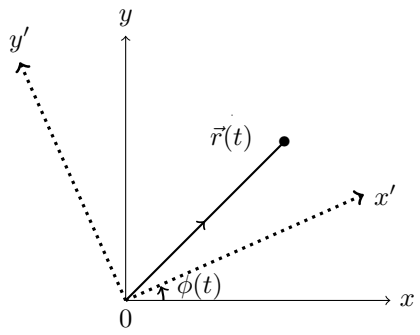

Figure 1.1: Rotating Frame

We denote the rotation by an angle ϕ, which parameterizes our unit vectors as such:

$$\begin{aligned} \vec{e}_r &= \cos\phi\vec{e}_x + \sin\phi\vec{e}_y \\ \vec{e}_\phi &= -\sin\phi\vec{e}_x + \cos\phi\vec{e}_y \end{aligned} \tag{1.4}$$

Note that $\vec{e}_r$ is a function of ϕ. So we can simply write an arbitrary position as $\vec{r} = r\vec{e}_r$. What is its derivative in time then? Well, we have to use the chain rule to get $\dot{\vec{r}} = \dot{r}\vec{e}_r + r\dot{\vec{e}}_r$ since time derivatives of unit vectors are no longer trivial. Using differential analysis, we can write this as

$$\begin{aligned} \dot{\vec{e}}_r &= \frac{\vec{e}_{r'} - \vec{e}_r}{dt} \\ &= \frac{(\cos(d\phi) - 1)\vec{e}_r - \sin(d\phi)\vec{e}_\phi}{dt} \\ &\approx \frac{\left((1 - \frac{d\phi^2}{2!}) - 1\right)\vec{e}_r - (d\phi - \frac{d\phi^3}{3!})\vec{e}_\phi}{dt} \\ &= \frac{d\phi}{dt}\vec{e}_\phi = \dot{\phi}\vec{e}_\phi \\ \Rightarrow\ &\boxed{\dot{\vec{e}}_r = \dot{\phi}\vec{e}_\phi} \end{aligned} \tag{1.5}$$

From this, we can thus write our **velocity** vector in radial coordinates as:

$$\boxed{\dot{\vec{r}} = \dot{r}\vec{e}_{\vec{r}} + r\dot{\phi}\vec{e}_\phi} \tag{1.6}$$

where we used the Taylor expansion above and ignored higher order terms. We can then do the

same for the **acceleration** (second derivative) to get:

$$\begin{aligned} \ddot{\vec{r}} &= \ddot{r}\vec{e}_r + \dot{r}\dot{\vec{e}}_r + \ddot{\phi}r\vec{e}_\phi + \dot{\phi}\dot{r}\vec{e}_\phi + \dot{\phi}r\dot{\vec{e}}_\phi \\ \Rightarrow \quad &\boxed{\ddot{\vec{r}} = (\ddot{r} - r\dot{\phi}^2)\vec{e}_r + (\ddot{\phi}r + 2\dot{r}\dot{\phi})\vec{e}_\phi} \end{aligned} \tag{1.7}$$

In equation (1.7) above, we have 4 second derivative terms which correspond, in order from left to right, to 1) *Linear*, 2) *centripetal*, 3) *angular* and 4) *Coriolis* accelerations. When we rotate an axis, it is also important to ask which axis we are rotating around and how fast we are rotating. In 2D these can be encompassed by one vector quantity $\vec{\Omega} = \dot{\phi}\vec{e}_z$. Here, $\dot{\phi}$ is the angular velocity and $\vec{e}_z$ is the axis of rotation. This gives a tool to elegantly derive the conjugate unit vector velocities.

- Along the $\vec{e}_r$ direction

$$\begin{aligned} \dot{\vec{e}}_r &= \vec{\Omega} \times \vec{e}_r \\ &= \dot{\phi}(\vec{e}_z \times \vec{e}_r) = \dot{\phi}\vec{e}_\phi \end{aligned} \tag{1.8}$$

- Along the $\vec{e}_\phi$ direction

$$\begin{aligned} \dot{\vec{e}}_\phi &= \vec{\Omega} \times \vec{e}_\phi \\ &= \dot{\phi}(\vec{e}_z \times \vec{e}_\phi) = \dot{\phi}(-\vec{e}_r) \end{aligned} \tag{1.9}$$

In general, considering a 3D case, the axis of rotation can be in a generic direction $\vec{e}_u$, i.e. $\vec{\Omega} = \Omega\vec{e}_u$ where Ω is the amplitude of the angular velocity. Given another generic unit vector $\vec{e}_v$, its time derivative is given by $\dot{\vec{e}}_v = \Omega \times \vec{e}_v$.

§1.3 Moving Local Frames

We now want to extend Newton's second law into a *local*, possibly time varying reference frame. To do this, we need to have a larger **global** fixed reference frame, from which we define relative coordinates to the smaller **local** reference frame. Let us denote $\vec{R}$ as the vector from the global origin to the local origin, and $\vec{\rho}$ as the local coordinate position. The global coordinate position is then $\vec{r} = \vec{R} + \vec{\rho}$. Also, we denote the global Cartesian coordinates with upper case X and Y, whereas the local Cartesian coordinates are denoted with lower case x and y.

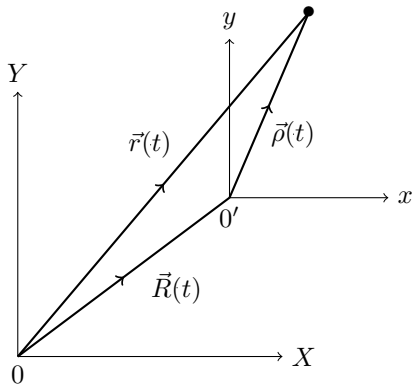

Figure 1.2: Moving Local Reference Frame w.r.t Global Frame

Let's first look at velocity $\dot{\vec{r}}$.

$$\begin{aligned}\dot{\vec{r}} &= \dot{\vec{R}} + \dot{\vec{\rho}} \\ \dot{\vec{R}} &= \dot{X}\vec{e}_x + \dot{Y}\vec{e}_y + \dot{Z}\vec{e}_z \\ \dot{\vec{\rho}} &= \left(\dot{x}\vec{e}_x + \dot{y}\vec{e}_y + \dot{z}\vec{e}_z\right) + \left(x\dot{\vec{e}}_x + x\dot{\vec{e}}_x + x\dot{\vec{e}}_x\right)\end{aligned} \tag{1.10}$$

We immediately see that we can use equationss (1.8) and (1.9) to simplify the $\dot{\vec{\rho}}$ vector as:

$$\begin{aligned}\dot{\vec{\rho}} &= \left(\dot{x}\vec{e}_x + \dot{y}\vec{e}_y + \dot{z}\vec{e}_z\right) + \left(x\dot{\vec{e}}_x + y\dot{\vec{e}}_y + z\dot{\vec{e}}_z\right) \\ &= \dot{\vec{\rho}}_r + \vec{\Omega} \times \vec{\rho}\end{aligned} \tag{1.11}$$

For which we have defined $\dot{\vec{\rho}}_r = \dot{x}\vec{e}_x + \dot{y}\vec{e}_y + \dot{z}\vec{e}_z$. Take note of this definition, it is **important** and is essentially the time derivative of $\vec{\rho}$ if we take the local frame as **stationary**. We also used the fact that

$$\begin{aligned}\dot{\vec{e}}_x &= \vec{\Omega} \times \vec{e}_x \\ \dot{\vec{e}}_y &= \vec{\Omega} \times \vec{e}_y \\ \dot{\vec{e}}_z &= \vec{\Omega} \times \vec{e}_z\end{aligned} \tag{1.12}$$

As a result, we have the following simple form for the velocity:

$$\boxed{\dot{\vec{r}} = \dot{\vec{R}} + \dot{\vec{\rho}}_r + \vec{\Omega} \times \vec{\rho}} \tag{1.13}$$

Now we move on to the acceleration $\ddot{\vec{r}} = \frac{d}{dt}\left(\dot{\vec{R}} + \dot{\vec{\rho}}_r + \vec{\Omega} \times \vec{\rho}\right)$. Again, we divide and conquer, evaluating each second derivative individually as follows.

$$\frac{d}{dt}\dot{\vec{R}} = \ddot{\vec{R}} = \ddot{X}\vec{e}_x + \ddot{Y}\vec{e}_y + \ddot{Z}\vec{e}_z \tag{1.14}$$

$$\frac{d}{dt}(\vec{\Omega} \times \vec{\rho}) = (\dot{\vec{\Omega}} \times \vec{\rho}) + (\vec{\Omega} \times \frac{d}{dt}\vec{\rho}) = (\dot{\vec{\Omega}} \times \vec{\rho}) + (\vec{\Omega} \times \dot{\vec{\rho}}_r) + \vec{\Omega} \times (\vec{\Omega} \times \vec{\rho}) \tag{1.15}$$

$$\frac{d}{dt}\dot{\vec{\rho}}_r = \ddot{\vec{\rho}}_r + \vec{\Omega} \times \dot{\vec{\rho}}_r \tag{1.16}$$

All these put together gives:

$$\boxed{\ddot{\vec{r}} = \ddot{\vec{R}} + \ddot{\vec{\rho}}_r + (\dot{\vec{\Omega}} \times \vec{\rho}) + \left(\vec{\Omega} \times (\vec{\Omega} \times \vec{\rho})\right) + (2\vec{\Omega} \times \dot{\vec{\rho}}_r)} \tag{1.17}$$

Where the 5 terms (from left to right) correspond to: 1) *local frame origin*, 2) *object w.r.t local frame*, 3) *angular component*, 4) *centripetal* and 5) *Coriolis* accelerations.

Chapter 2

Systems of Particles and Rigid Bodies

In this chapter, we will be looking at the dynamics of systems of many particles and rigid bodies. We will be tackling these in sequence so let us first consider a system of many particles where each particle is labelled by an index i. As such, the position of each particle is given by $\vec{r}_i$, whereby they can experience internal forces $\vec{f}_i$ and external forces $\vec{F}_i$. As such, we have by Newton's second law that:

$$m_i\ddot{\vec{r}}_i = \vec{F}_i + \vec{f}_i \tag{2.1}$$

This may seem like a trivial restatement of Newton's law but it would be good to internalize this as we progress along the chapter.

§2.1 Center of Mass Motion

We now zoom out to look at the entirety of the system and sum over all the forces on the individual particles. As such, we get:

$$\sum_i m_i\ddot{\vec{r}}_i = \sum_i(\vec{F}_i + \vec{f}_i) = \sum_i \vec{F}_i \equiv \vec{F} \tag{2.2}$$

where we utilized Newton's third law which says that for every internal force $\vec{f}_i$, there will be an equal and opposite force such that the sum over all internal forces will result in a trivial net force. As such, we can simply look at the *center of mass* of this system for much of our analysis. The center of mass of a system of particles is defined by:

$$\vec{r}_c = \frac{\sum_i m_i\vec{r}_i}{\sum_i m_i} = \frac{\sum_i m_i\vec{r}_i}{M} \tag{2.3}$$

$$\Rightarrow \boxed{M\ddot{\vec{r}}_c = \vec{F}} \tag{2.4}$$

§2.2 Kinetic Energy of a System of Particles

The total *kinetic energy* E_k of a system of particles is the sum of all the individual particle kinetic energies.

$$E_k = \frac{1}{2}\sum_i m_i v_i^2 \tag{2.5}$$

To simplify our analysis, we define a vector $\vec{\rho}_i = \vec{r}_i - \vec{r}_c$ (the position of particle i from the system center of mass). From this, we have:

$$\vec{r}_i = \vec{r}_c + \vec{\rho}_i \tag{2.6}$$
$$\Rightarrow\ \dot{\vec{r}}_i = \dot{\vec{r}}_c + \dot{\vec{\rho}}_i \tag{2.7}$$
$$\Rightarrow\ v_i^2 = \dot{\vec{r}}_c \cdot \dot{\vec{r}}_c + \dot{\vec{\rho}}_i \cdot \dot{\vec{\rho}}_i + 2\dot{\vec{r}}_c \cdot \dot{\vec{\rho}}_i \tag{2.8}$$
$$\Rightarrow\ E_k = \frac{1}{2}\sum_i m_i \dot{\vec{r}}_c \cdot \dot{\vec{r}}_c + \frac{1}{2}\sum_i m_i \dot{\vec{\rho}}_i \cdot \dot{\vec{\rho}}_i + \sum_i m_i (\dot{\vec{r}}_c \cdot \dot{\vec{\rho}}_i) \tag{2.9}$$

Looking at the third term in the sum above, we notice that

$$\begin{aligned}\sum_i m_i(\dot{\vec{r}}_c \cdot \dot{\vec{\rho}}_i) &= \dot{\vec{r}}_c \cdot \left(\sum_i m_i \dot{\vec{\rho}}_i\right) \\ &= \dot{\vec{r}}_c \cdot \frac{d}{dt}\left(\sum_i m_i \vec{\rho}_i\right) = 0\end{aligned} \tag{2.10}$$

The vanishing result is because $\vec{\rho}_i$ is the relative distance to the center of mass, for which by the definition of the center of mass, the sum over all these relative distances weighted by the particle masses will remain time-invariant. Thus, the kinetic energy simplifies to:

$$\boxed{E_k = \frac{1}{2}\sum_i m_i\ \dot{\vec{r}}_c \cdot \dot{\vec{r}}_c + \frac{1}{2}\sum_i m_i\ \dot{\vec{\rho}}_i \cdot \dot{\vec{\rho}}_i} \tag{2.11}$$

Where the 2 terms (from left to right) correspond to: 1) *the kinetic energy of the center of mass* 2) *the kinetic energy relative to the center of mass.*

§2.3 Angular Momentum

The angular momentum $\vec{H}$ of a single particle at some position $\vec{r}$ is given by:

$$\vec{H} = \vec{r} \times \vec{p} \tag{2.12}$$

Now bringing this to the context of many-body systems in a fixed reference frame, we simply index the above equation for the i-th particle. The total angular momentum of the system $\vec{H}$ is then:

$$\vec{H} = \sum_i \vec{r}_i \times m_i \dot{\vec{r}}_i \tag{2.13}$$

Let's now take the time derivative of this. We get:

$$\begin{aligned}\dot{\vec{H}} &= \sum_i (\dot{\vec{r}}_i \times m_i \dot{\vec{r}}_i) + \sum_i (\vec{r}_i \times m_i \ddot{\vec{r}}_i) \\ &= \sum_i \vec{r}_i \times m_i \ddot{\vec{r}}_i \\ &= \sum_i \vec{r}_i \times (\vec{F}_i + \vec{f}_i) \\ &= \sum_i \vec{r}_i \times \vec{F}_i = \vec{M}\end{aligned} \tag{2.14}$$

where $\vec{M}$ is known as the *momenta of external forces* acting on the system. Make sure how we get from each step to the next above. To restate the above result:

$$\boxed{\dot{\vec{H}} = \vec{M}} \tag{2.15}$$

This all seems reasonable, but let's now consider again a fixed reference frame and some point P ($\vec{r}_P$) that is **not** the center of mass. We also have a particle labelled with i situated away from P, with a relative position $\vec{s}_i$. The position of the particle from the origin O is then $\vec{r}_i = \vec{r}_P + \vec{s}_i$. If the particle is moving with velocity $\dot{\vec{r}}_i$, its angular momentum with respect to P is:

$$\vec{H}_P = \sum_i \vec{s}_i \times m_i \dot{\vec{r}}_i \tag{2.16}$$

$$\Rightarrow \quad \dot{\vec{H}}_P = \sum_i \vec{s}_i \times m_i \ddot{\vec{r}}_i + \dot{\vec{s}}_i \times m_i \dot{\vec{r}}_i \tag{2.17}$$

We note that $\dot{\vec{s}}_i = \dot{\vec{r}}_i - \dot{\vec{r}}_P$ and $\dot{\vec{r}}_i = \dot{\vec{r}}_c + \dot{\vec{\rho}}_i$ where $\vec{r}_c$ is the position to the center of mass and $\vec{\rho}_i$ is defined relative to that. So, the second term in the derivative of angular momentum above becomes:

$$\begin{aligned}&\sum_i (\dot{\vec{r}}_i - \dot{\vec{r}}_P) \times m_i (\dot{\vec{r}}_c + \dot{\vec{\rho}}_i) = -\sum_i \dot{\vec{r}}_P \times m_i \dot{\vec{r}}_c = -\dot{\vec{r}}_P \times \dot{\vec{r}}_c M \\ \Rightarrow \quad &\dot{\vec{H}}_P = \sum_i \vec{s}_i \times m_i \ddot{\vec{r}}_i - (\dot{\vec{r}}_P \times \dot{\vec{r}}_c) M \\ \Rightarrow \quad &\boxed{\dot{\vec{H}}_P = \vec{M}_P - (\dot{\vec{r}}_P \times \dot{\vec{r}}_c) M}\end{aligned} \tag{2.18}$$

where we have defined $M \equiv \sum_i m_i$. So this (equation 2.18) is the more general version of equation 2.15, where equality between these 2 equations hold only if:

- The position of the point P is itself at the center of mass.
- All particles are stationary ($\dot{\vec{r}}_P$ or $\dot{\vec{r}}_c$).
- All particles have velocity parallel to $\dot{\vec{r}}_c$.

§2.4 Rigid Body kinematics

Definition 2.4.1. Rigid Body: *A system of particles for which their relative distances $\vec{\rho}_{ij}$ are fixed in magnitude.*

$$\|\vec{\rho}_{ij}\| = \text{constant} \tag{2.19}$$

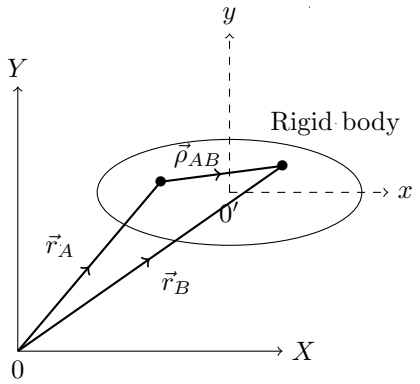

Figure 2.1: 2 Points A and B on a Rigid Body

Consider 2 points on a rigid body A and B which a relative position vector between them $\vec{\rho}_{AB} = \vec{r}_B - \vec{r}_A$. Let's also say that the rigid body is rotating with $\vec{\omega}$ with respect to it's local reference frame (denoted with axes x and y). Then we have:

$$\begin{aligned} \dot{\vec{r}}_B &= \dot{\vec{r}}_A + \dot{\vec{\rho}}_{AB} + \vec{\omega} \times \vec{\rho}_{AB} \\ &= \dot{\vec{r}}_A + \vec{\omega} \times \vec{\rho}_{AB} \end{aligned} \tag{2.20}$$

where the second term drops out because there is no relative motion between the points following the rigid body definition (2.4). The acceleration is then:

$$\begin{aligned} \ddot{\vec{r}}_{AB} &= \ddot{\vec{r}}_A + \ddot{\vec{\rho}}_{AB} + (\vec{\omega} \times \vec{\omega} \times \vec{\rho}_{AB}) + (\dot{\vec{\omega}} \times \vec{\rho}_{AB}) + 2(\vec{\omega} \times \dot{\vec{\rho}}_{AB}) \\ &= \ddot{\vec{r}}_A + (\vec{\omega} \times \vec{\omega} \times \vec{\rho}_{AB}) + (\dot{\vec{\omega}} \times \vec{\rho}_{AB}) \end{aligned} \tag{2.21}$$

So in **summary**, we get the 2 dynamical equations that summarize rigid body kinematics as:

$$\boxed{\dot{\vec{r}}_B = \dot{\vec{r}}_A + \vec{\omega} \times \vec{\rho}_{AB}} \tag{2.22}$$

$$\boxed{\ddot{\vec{r}}_{AB} = \ddot{\vec{r}}_A + (\vec{\omega} \times \vec{\omega} \times \vec{\rho}_{AB}) + (\dot{\vec{\omega}} \times \vec{\rho}_{AB})} \tag{2.23}$$

§2.4.1 Moment of Inertia

Earlier we saw that if we have any of the 3 conditions 1) the point is question is at the center of mass or 2) we are considering fixed points or 3) the particles have velocity parallel to $\dot{\vec{r}}_c$, this

implies $\dot{\vec{H}}_0 = \vec{M}_0$ where the subscript 0 indicates a specific particle/point. For an entire rigid body system of particles, we can get the angular momentum as follows:

$$\begin{aligned}\vec{H} &= \int \rho(\vec{r}_c \times \vec{v})dv \\ &= \int \rho\big(\vec{r}_c \times (\dot{\vec{R}}_c + \vec{\omega} \times \vec{r}_c)\big)dv \\ &= \int \rho(\vec{r}_c \times \vec{\omega} \times \vec{r}_c)dv\end{aligned} \tag{2.24}$$

Where $\vec{R}_c$ is the center of mass position and $\vec{r}_c$ is the position of an arbitrary point on the rigid body relative to the center of mass. Think about why some of the terms above have vanished allowing for the final simplification. Taking a closer look at the integrand:

$$\begin{aligned}\vec{r}_c \times \vec{\omega} \times \vec{r}_c &= \vec{r}_c \times \big((z\omega_y - y\omega_z)\vec{e}_x + (x\omega_z - z\omega_x)\vec{e}_y + (y\omega_x - x\omega_y))\vec{e}_z\big) \\ &= \big(+ \omega_x(y^2 + z^2) - \omega_y(xy) - \omega_z(xz)\big)\vec{e}_x \\ &+ \big(- \omega_x(yx) + \omega_y(x^2 + y^2) - \omega_z(yz)\big)\vec{e}_y \\ &+ \big(- \omega_x(zx) - \omega_y(zy) + \omega_z(x^2 + y^2)\big)\vec{e}_z\end{aligned} \tag{2.25}$$

Already, we see some remnants of a matrix structure starting to take shape, so let us define the following *real*, *symmetric* matrix:

$$\hat{I} = \begin{bmatrix} I_x & -I_{xy} & -I_{xz} \\ -I_{xy} & I_y & -I_{yz} \\ -I_{xz} & -I_{yz} & I_z \end{bmatrix} \tag{2.26}$$

where the matrix components are defined as follows:

$$\begin{aligned} I_x &= \int \rho(y^2 + z^2)dV, \quad I_y = \int \rho(z^2 + x^2)dV, \quad I_z = \int \rho(x^2 + y^2)dV \\ I_{xy} &= \int \rho(x \cdot y)dV, \quad I_{xz} = \int \rho(x \cdot z)dV, \quad I_{yz} = \int \rho(y \cdot z)dV \end{aligned} \tag{2.27}$$

$\hat{I}$ is often referred to as the *inertia tensor* (matrices are 2-rank tensors). From this, we get:

$$\boxed{\vec{H} = \hat{I}\vec{\omega}} \tag{2.28}$$

where $\vec{\omega} = \omega_x\vec{e}_x + \omega_y\vec{e}_y + \omega_z\vec{e}_z$. What if we want to compute the moment of inertia for an object that has a more complex shape? We can utilize 3 properties of the moment of inertia to aid us in this. These properties are:

1. **Parallel Axis Theorem**:

Theorem 2.4.1. Parallel Axis Theorem: *Given a system's moment of inertia about an axis through the object's center of gravity (unprimed) and the relative position $\vec{R}$ from center of mass axis to a parallel axis (primed), we have the moment of inertia about the primed axis as:*

$$I'_{ij} = I_{ij} + M(\left\|\vec{R}\right\|^2 \delta_{ij} - R_i R_j) \tag{2.29}$$

where M is the rigid body mass and δ_{ij} is the Kronecker-delta function.

2. $\hat{I}$ **in Rotated Coordinates**:

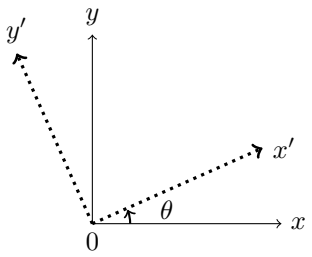

Figure 2.2: Rotated Coordinates

As show in the figure above, if we consider the rotated primed coordinates by an angle θ about the z-axis, we can represent this rotation by the *rotation matrix*:

$$R = \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{2.30}$$

Of course, this is **not** the general form of a rotation matrix one can construct. The moment of inertia in these rotated coordinates is then given by:

$$\boxed{\hat{I}' = R\hat{I}R^T} \tag{2.31}$$

Note: The reason why we have to left **and** right multiply a rotation matrix is because $\hat{I}$ is a 2-rank tensor (has 2 indices), so we have to transform both indices (as opposed to just once like for vectors with 1 index).

The axes whereby the angular momentum tensor is strictly **diagonal** are as the *principal axes.*

3. **Linearity of Moment of Inertia** $\hat{I}_{\text{total}} = \sum_j \hat{I}_j$:

This is simply saying that the total moment of inertia of a rigid body is equal to the sum of the moments of inertia of its constituent parts.

§2.4.2 Kinetic Energy of Rigid Bodies

Consider some rigid body with a differential element of volume dV moving with constant velocity $\vec{v}$, the kinetic energy of this element is given by:

$$E_{k,dV} = \frac{1}{2}\rho(\vec{v}\cdot\vec{v})dV \tag{2.32}$$

As such, the kinetic energy of the entire rigid body would be the of all the kinetic energy contributions from all differential volume elements:

$$\begin{aligned} E_k &= \frac{1}{2}\int \rho(\vec{v}\cdot\vec{v})dV \\ &= \frac{1}{2}\int \rho(\vec{v}_c+\vec{\omega}\times\vec{r}_c)\cdot(\vec{v}_c+\vec{\omega}\times\vec{r}_c)dV \\ &= \frac{1}{2}\int \rho\|\vec{v}_c\|^2 dV + \int \rho\vec{v}_c\cdot(\vec{\omega}\times\vec{r}_c)dV + \frac{1}{2}\int \rho(\vec{\omega}\times\vec{r}_c)\cdot(\vec{\omega}\times\vec{r}_c)dV \\ &= \frac{1}{2}M\|\vec{v}_c\|^2 + \frac{\vec{\omega}}{2}\cdot\int \rho\big(\vec{r}_c\times(\vec{\omega}\times\vec{r}_c)\big)dV \\ &= \frac{1}{2}Mv_c^2 + \frac{1}{2}\vec{\omega}\cdot\vec{H}_c \end{aligned} \tag{2.33}$$

In our analysis above, there are a few important things to take note of. First, we saw that $\int \rho\vec{v}_c\cdot(\vec{\omega}\times\vec{r}_c)dV = 0$. This is due to the fact that $\vec{v}_c$ and $\vec{\omega}$ are constant, making the average over velocities relative to the center of mass vanish. Also, we utilized the fact that $\vec{H}_c = \vec{r}_c\times(\vec{\omega}\times\vec{r}_c)$ to simplify our answer in the final step. As such, we restate the elegant result just derived again:

$$\boxed{E_k = \frac{1}{2}Mv_c^2 + \frac{1}{2}\vec{\omega}\cdot\vec{H}_c} \tag{2.34}$$

Recall that $M = \sum_i m_i$.

Chapter 3

Analytical Mechanics

*We are now moving away from the **Newtonian** approach to mechanics to a new formalism of studying dynamical systems. This is known as the **Lagrangian formalism** which is a reformulation of classical mechanics by Joseph-Louis Lagrange in 1788. We will be introducing this new means of analysis via an example.*

§3.1 Newton vs Lagrange

Consider a system pendulum system as illustrated in figure 3.1. We will work through this problem using both the Newtonian and Lagrangian means of analysis to show the differences in application and to hopefully give some motivation as to why we would want to adopt Lagrange's approach.

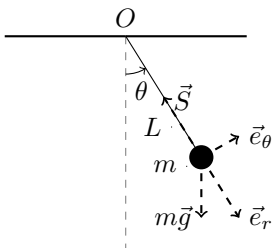

Figure 3.1: Simple Pendulum System

§3.1.1 The Newtonian Method

We will utilize polar coordinates (r, θ) here for convenience to describe our system. By Newton's second law, we have:

$$\begin{aligned} m\ddot{\vec{r}} &= \sum \vec{F} \\ m\big((\ddot{r} - r\dot{\theta}^2)\vec{e}_r + (r\ddot{\theta} + 2\dot{r}\dot{\theta})\vec{e}_\theta\big) &= m\vec{g} + \vec{S} \end{aligned} \tag{3.1}$$

where $\vec{S}$ is the so called *constraint force*. It is called as such because it constrains the mass to only swing along the circumference of the circle with radius L. Now if we decompose the forces acting on the mass into the $\vec{e}_r$ and $\vec{e}_\theta$ components, we have the following 2 equations:

$$m(\ddot{r} - r\dot{\theta}^2) = mg\cos\theta - S \tag{3.2}$$

$$m(r\ddot{\theta} + 2\dot{r}\dot{\theta}) = -mg\sin\theta \tag{3.3}$$

As seen above, the first equation defines the constraint force whereas the second can be used to find the equation of motion of the mass. Now, we enforce the constraint that is inherently applied on this system. We have that $r = L \;\Rightarrow\; \dot{r} = \ddot{r} = 0$. Application of this constraint leads to:

$$\ddot{\theta} = -\frac{g}{L}\sin\theta \tag{3.4}$$

§3.1.2 The Lagrangian Method

In Lagrangian mechanics, the way we approach/think about a problem is slightly different. The first step is to identify the *number of degrees of freedom* of the system.

Definition 3.1.1. Degrees of Freedom: *For a given mechanical system, its degrees of freedom are the number of independent parameters required to uniquely define its configuration.*

For this example, we see that we only require one dynamical variable θ to comprehensively describe the relevant dynamics. So the number of degrees of freedom here is simply 1. Furthermore, we term any such dynamical variable that specifies the system configuration as a *generalized coordinate* (Often denoted as q_j). Right now, we are simply going to state the *Euler-Lagrange equation* which is the general form for a system's equation of motion given its *Lagrangian* $\mathcal{L}(q_j, \dot{q}_j)$.

$$\boxed{\frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{q}_j} - \frac{\partial \mathcal{L}}{\partial q_j} = 0} \tag{3.5}$$

Where the Lagrangian of a system is defined as:

$$\boxed{\mathcal{L}(q_j, \dot{q}_j) = T(q_j, \dot{q}_j) - V(q_j)} \tag{3.6}$$

Where $T(q_j, \dot{q}_j)$ is the kinetic energy of the system and $V(q_j)$ is the potential energy.

Note: The lagrangian is **not** the total energy of the system since there is a negative sign attached to the potential energy $V(q_j)$.

So going back to our pendulum system, we only have one generalized coordinate $q_1 = \theta$. So our Lagrangian can be written as follows:

$$\mathcal{L} = \frac{1}{2}m(L\dot{\theta})^2 - (-mg\cos\theta) = \frac{1}{2}mL^2\dot{\theta}^2 + mg\cos\theta \tag{3.7}$$

Substituting this into equation 3.5, we get:

$$\begin{aligned} &\frac{d}{dt}(mL^2\dot{\theta}) + mgL\sin\theta = 0 \\ \Rightarrow\ &\ddot{\theta} = -\frac{g}{L}\sin\theta \end{aligned} \tag{3.8}$$

which is exactly the equation of motion we derived from the Newtonian approach (just much faster). So we see that in analytical mechanics, instead of computing accelerations and working with constraint forces, we simply define generalized coordinates and work with energy functions to derive the equations of motion. All accelerations are implicitly computed when we apply the Euler-Lagrange equation. Amazing!

§3.2 Virtual Displacement and Work

We are now going to explore the concepts of *virtual displacement* and *virtual work* which will then lead to a derivation of the Euler-Lagrange equation (3.5). First, let us consider a rigid body that moves in 3D space with respect to some origin O. We pick some point P on this body and denote the position of this point with respect to the origin as $\vec{r}$. In general, this vector is a function of time and spatial coordinates $\vec{r} = \vec{r}(t, x, y, z)$.

Additionally, we have that this body is subject to a constraint (fixture) such that the point on the body attached to the fixture (not point P) cannot move but the body can rotate about that point. If we apply an external force on this body, it would move to a new position such that point P would move from $\vec{r}$ to $\vec{r}'$ by $d\vec{r} = \vec{r}' - \vec{r}$. We then use the product rule to expand out this differential:

$$d\vec{r} = \frac{\partial\vec{r}}{\partial dt}dt + \frac{\partial\vec{r}}{\partial x}dx + \frac{\partial\vec{r}}{\partial y}dy + \frac{\partial\vec{r}}{\partial z}dz \tag{3.9}$$

We call this the *actual displacement* and is consistent with the constraints imposed on the body along with the differential equation of motion.

Similarly, we can define a *virtual displacement* of the body. To do this again consider the system we described above. The main difference between the virtual and actual displacements is that although it is consistent with the system's constraints, it does **not** satisfy the equation of motion (i.e. only a displacement with respect to the spatial coordinates, time is held fixed).

$$\delta\vec{r} = \vec{r}'' - \vec{r} = \frac{\partial\vec{r}}{\partial x}\delta x + \frac{\partial\vec{r}}{\partial y}\delta y + \frac{\partial\vec{r}}{\partial z}\delta z \tag{3.10}$$

The idea here is that virtual displacement is any possible displacement (or variation in path) that adheres to the constraints. So it is as though we are taking a snapshot in time and observing all available (but not necessarily physically realizable) paths our particle can take. As an extension of this idea, we can also define a *virtual work* from this virtual displacement. First recall the conventional definition of work:

Definition 3.2.1. *The work done by a force $\vec{F}$ is the path integral of its scalar tangential component along the path at its point of application.*

$$W = \int_a^b \vec{F} \cdot d\vec{s} \tag{3.11}$$

So drawing from this intuition, the virtual work of a force $\vec{F}$ over a virtual displacement $\delta\vec{r}$ is given by:

$$\boxed{\delta W = \vec{F} \cdot \delta\vec{r}} \tag{3.12}$$

Note: If the direction of the force and virtual displacements are orthogonal, then the virtual work is 0 by definition.

§3.3 Constraints and Constraint Forces

Consider 2 particles that move in 3D space while being connected by a rigid rod of fixed length L. We shall denote the positions of the particles by $\vec{r}_1$ and $\vec{r}_2$. The rigid rod imposes a *kinematic constraint* which we can write as $(\vec{r}_2 - \vec{r}_1) \cdot (\vec{r}_2 - \vec{r}_1) = L^2$. In a more general sense, constraints that can be written in the form:

$$f(t, \vec{r}_1, \vec{r}_2, ...) = 0 \tag{3.13}$$

are called *holonomic constraints*.

Definition 3.3.1. Holonomic Constraints: *For a given physical system, holonomic constraints are constraints that:*

1. Are **equality** constraints.
2. Depend **only on the configuration** of the system and **not the velocities**.

In this class, we will only consider holonomic constraints. A simple example of a holonomically constrained system is a slider that moves along a circular guide of radius R in the x, y-plane.

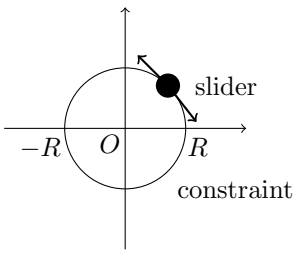

Figure 3.2: Slider Constrained to a Circle

The motion of the slider can be described by 2 coordinates (x and y) but there exists a non-trivial relation between them (holonomic constraint) which can be written as $x^2 + y^2 = R^2$. On the other hand, an example of a non-holonomic constraint is a particle that is constrained to move along or inside the circle, which renders the constraint being $x^2 + y^2 \leq R^2$. The point of formalizing this is because **only** holonomic constraints give rise to constraint forces $\vec{F}^c$. As such, we can decompose constraint forces into 2 components:

$$\vec{F}^c = \vec{F}^c_{ideal} + \vec{F}^c_{non-ideal} \tag{3.14}$$

where the ideal constraint forces $\vec{F}^c_{ideal}$ are forces **perpendicular** to virtual displacement. As such, the component of virtual work done by $\vec{F}^c_{ideal}$ will always vanish (equate to 0).

$$\boxed{\delta W_{\vec{F}^c_{ideal}} = \vec{F}^c_{ideal} \cdot \delta\vec{r} = 0} \tag{3.15}$$

§3.4 Generalized Coordinates

We will now look at the concept of *generalized coordinates* which will allow us to explore parameterization of the configuration of a mechanical system.

Definition 3.4.1. Generalized Coordinates: *A non-unique set of independent coordinates that uniquely represent the configuration of a system at **every** moment of time.*

The number of generalized coordinates is the **same** as the number of degrees of freedom (3.1.2) of the system in question. In general (as mentioned previously), we denote these generalized coordinates with q_j. So for instance, if some system has n degrees of freedom, then we can introduce n generalized coordinates:

$$\vec{r}_i = \vec{r}_i(t, q_1, q_2, ..., q_n) \tag{3.16}$$

There is a simple but powerful relation for the number of degrees of freedom that will be essential in solving any mechanics problem via the Lagrangian approach.

$$\boxed{\text{(No. constrained DoF)} = \text{(No. unconstrained DoF)} - \text{(No. of constraints)}} \tag{3.17}$$

where DoF stands for degrees of freedom. Let's consider 2 simple examples to make sure we know how to apply this concept.

Example 1:

Consider again the slider contrained to move along a circle as illustrated in figure 3.2. The unconstrained coordinates which uniquely describe the slider would be x and y since we are only considering motion in the x, y-plane. However, x and y are related by constraint relation, making them **not** independent and hence **not** generalized coordinates. So let us now apply our formula 3.4. We have:

1. 2 unconstrained degrees of freedom (x and y).

2. 1 holonomic constraint ($x^2 + y^2 = R^2$).

As such, the *number of constrained degrees of freedom* $= 2 - 1 = 1$. This means we can simply introduce just 1 generalized coordinate (e.g. θ, $s = R\theta$, etc) that will allow us to fully describe the dynamics of our system.

Example 2:

Now consider instead a rigid wheel constrained to spin around the z-axis in the x, y-plane. This is illustrated in figure 3.3. Rigid bodies in 3D have 6 degrees of freedom, 3 of which are translational and the remaining 3 are rotational. Our set-up is such that we have imposed 3 constraints:

1. 2 rotational constraints since we only permit rotation along the z-axis.
2. 1 translational since we only permit translation in the x, y-plane.

As such, we get that the *number of constrained degrees of freedom* $= 6 - 3 = 3$, which for instance could have the generalized coordinates x, y and θ.

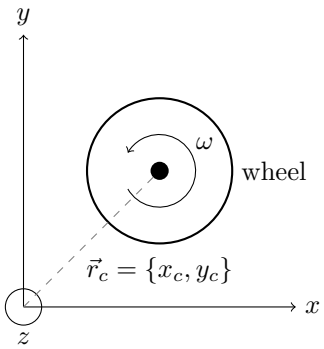

Figure 3.3: Rigid Wheel Rotating in the x, y-Plane

§3.5 D'ALembert's Principle

Previously, we have seen that the Euler-Lagrange equation does not depend on the constraint forces. However, it is still unclear on how we were allowed to do this. *D'Alembert's principle* gives us an answer to this question. Let us consider a simple system of a point mass m with position denoted by $\vec{r}$ and subject to a path constraint. The forces acting on this point mass are 1) gravity (mg) 2) constraint forces. Let's first try to do a little massaging to Newton's second law:

$$\begin{aligned} m\ddot{\vec{r}} &= \sum \vec{F} \\ \Rightarrow \quad \sum \vec{F} + (-m\ddot{\vec{r}}) &= 0 \end{aligned} \tag{3.18}$$

We now define $\vec{F}^I = -m\ddot{\vec{r}}$ as the *inertial force*. The equation above is also often referred to as the *equation of dynamic equilibrium*. Thus for our simple system in question, we have:

$$\sum \vec{F} + \vec{F}^I = m\vec{g} + \vec{F}^c_{ideal} + \vec{F}^c_{non-ideal} + \vec{F}^I = 0 \tag{3.19}$$

Recalling our definition of virtual work, if we instead consider the virtual work done on the point mass, we get:

$$\begin{aligned}\delta W &= \left(m\vec{g} + \vec{F}^c_{ideal} + \vec{F}^c_{non-ideal} + \vec{F}^I\right) \cdot \delta\vec{r} \\ &= \left(m\vec{g} + \vec{F}^c_{non-ideal} + \vec{F}^I\right) \cdot \delta\vec{r} = 0\end{aligned} \tag{3.20}$$

where the $\vec{F}^c_{ideal}$ has dropped out due to orthogonality and we are left with a **scalar** condition for dynamic equilibrium. So in general, we get the relation:

$$\left(\sum \vec{F}_{non-constraint} + \vec{F}^c_{non-ideal}\right) \cdot \delta\vec{r} + \left(\vec{F}^I\right) \cdot \delta\vec{r} = 0 \tag{3.21}$$

The result above is known as the *D'Alembert's principle.*

D'Alembert's Principle:

The sum of virtual work done by non-constraint and non-ideal constraint forces and the virtual work done by inertial forces must vanish.

$$\delta W + \delta W^I = 0 \tag{3.22}$$

This principle is key to deriving the Euler-Lagrange equations of motion.

§3.6 The Euler-Lagrange Equation

Going back to our analysis of the single particle along a constrained trajectory, it gave us 2 components in the result, $\left(\vec{F} + \vec{F}^c_{non-ideal}\right) \cdot \delta\vec{r}$ and $\left(\vec{F}^I\right) \cdot \delta\vec{r}$ where are just denoting all non-constraint forces as $\vec{F}$. Since the particle is moving along this guided path, it has 1 degree of freedom, implying only one generalized coordinate. Hence, we have that $\vec{r} = \vec{r}(q)$. So we have that:

$$\delta\vec{r} = \frac{\partial\vec{r}}{\partial q}\delta q \tag{3.23}$$

$$\Rightarrow \quad \delta W = \left(\vec{F} + \vec{F}^c_{non-ideal}\right) \cdot \delta\vec{r} = \left(\vec{F} + \vec{F}^c_{non-ideal}\right) \cdot \left(\frac{\partial\vec{r}}{\partial q}\delta q\right)$$

$$\Rightarrow \quad \delta W = \left(\left(\vec{F} + \vec{F}^c_{non-ideal}\right) \cdot \frac{\partial\vec{r}}{\partial q}\right)\delta q \equiv Q_q \delta q \tag{3.24}$$

So we have this new quantity Q_q which we will call the *generalized force.* On the other hand, we still have the virtual inertial work, from which we can do a similar derivation from that above

and get the *generalized inertial force* Q_q^I:

$$\begin{aligned} Q_q^I &= -m\ddot{\vec{r}}\cdot\frac{\partial\vec{r}}{\partial q} \\ &= -m\Big[\frac{d}{dt}\Big(\dot{\vec{r}}\frac{\partial\vec{r}}{\partial q}\Big) - \dot{\vec{r}}\frac{d}{dt}\Big(\frac{\partial\vec{r}}{\partial q}\Big)\Big] \\ &= -m\Big[\frac{d}{dt}\Big(\dot{\vec{r}}\frac{\partial\dot{\vec{r}}}{\partial\dot{q}}\Big) - \dot{\vec{r}}\Big(\frac{\partial\dot{\vec{r}}}{\partial q}\Big)\Big] \\ &= -\frac{d}{dt}\frac{\partial}{\partial\dot{q}}\Big[\frac{1}{2}m\dot{\vec{r}}\cdot\dot{\vec{r}}\Big] + \frac{\partial}{\partial q}\Big[\frac{1}{2}m\dot{\vec{r}}\cdot\dot{\vec{r}}\Big] \\ &= \Big[-\frac{d}{dt}\frac{\partial}{\partial\dot{q}} + \frac{\partial}{\partial q}\Big]T \end{aligned} \tag{3.25}$$

where T is the kinetic energy of the system. So we see that we can immediately compute the generalized inertial forces from just knowing the kinetic energy of the system! Now going back to D'Alembert's princple, we require that:

$$\begin{aligned} & \delta W + \delta W^I = 0 \\ \Rightarrow\ & (Q_q + Q_q^I)\delta q = 0 \\ \Rightarrow\ & Q_q + Q_q^I = 0 \\ \Rightarrow\ & \boxed{\frac{d}{dt}\frac{\partial T}{\partial\dot{q}} - \frac{\partial T}{\partial q} = Q_q} \end{aligned} \tag{3.26}$$

It is then common practice that we split the generalized force into *conservative* and *non-conservative* components $Q_q = Q_q^c + Q_q^{nc}$, the reason being that conservative forces can be easily derived knowing the potential energy function of the system. Thus, for a conservative system ($Q_q = Q_q^c$), we have that:

$$\frac{d}{dt}\frac{\partial T}{\partial\dot{q}} - \frac{\partial T}{\partial q} = Q_q^c \tag{3.27}$$

We also know that:

$$\delta W^c = -\delta V = -\frac{\partial V}{\partial q}\delta q = Q_q^c\delta q \tag{3.28}$$

where V is the potential energy of the system. These combine to give:

$$\frac{d}{dt}\frac{\partial T}{\partial\dot{q}} - \frac{\partial T}{\partial q} = -\frac{\partial V}{\partial q} \tag{3.29}$$

Note: The potential energy function **cannot** depend on generalized velocities ($-\frac{\partial V}{\partial\dot{q}} = 0$).

Keeping the above note in mind, we can rearrange all these terms to give:

$$\frac{d}{dt}\frac{\partial(T-V)}{\partial\dot{q}} - \frac{\partial(T-V)}{\partial q} = 0 \tag{3.30}$$

Finally, we define the *Lagrangian* of a system as $\mathcal{L} = T - V$, which results in the Euler-Lagrange equation for a system with a single degree of freedom:

$$\boxed{\frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{q}} - \frac{\partial \mathcal{L}}{\partial q} = 0} \tag{3.31}$$

Note: The Euler-Lagrange equation above (3.6) is valid **only** for conservative systems. For non-conservative systems, we need to utilize the more general form in equation 3.6.

For a system with more than 1 degree of freedom, we simply add an index to our generalized coordinates to get the previously presented equation (3.5). With this, we can then have a systematic work-flow to employ this new tool in tackling dynamics problems.

Lagrangian Work-Flow:

1. Identify the number of degrees of freedom of the system n.
2. Introduce the generalized coordinates $\{q_1, q_2, ..., q_n\}$.
3. Compute the kinetic energy:
$$T = T(t, q_1, ..., q_n)$$
.
4. Compute all conservative forces:
$$Q_j^c = -\frac{\partial V(q_1, ..., q_n)}{\partial q_j}$$
5. Compute all non-conservative forces:
$$\delta W^{nc} = \sum_{j=1}^{n} Q_j^{nc} \delta q_j$$
6. Form the Euler-Lagrange equation of motion:
$$\frac{d}{dt}\frac{\partial T}{\partial \dot{q}_j} - \frac{\partial T}{\partial q_j} = Q_j^c + Q_j^{nc}$$

Note: If we are considering rigid body mechanics, we can separate the kinetic energy term into 2 components: 1) kinetic energy due to translation of the center of mass, $T_{trans,CoM}$ 2) kinetic energy due to rotation about the center of mass, $T_{rot,CoM}$.

The Lagrangian satisfies a few neat properties. These are listed below.

Lagrangian Properties:

1. The Lagrangian is a **scalar** function. This follows from the nature of which it is defined.
2. For a system with only conservative forces, the Lagrangian scaled by a constant factor α gives us the same equations of motions and thus the same physics. The constant will drop out in the Euler-Lagrange equation.
3. Adding a scalar constant to the Lagrangian itself gives us the same equations of motion. $\mathcal{L} \sim \mathcal{L}' = \mathcal{L} + \alpha$
4. If we define $\mathcal{L}' = \mathcal{L} + \frac{d}{dt} f(q,t)$, this does not change the physics either since

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

   and adding constants do not change the *action* minimization problem.

Chapter 4

Conservation Laws

This week, we will look at conservation laws within the Lagrangian framework. Simply speaking, conserved quantities are quantities that remain invariant under time evolution. Specifically, we will be looking at the conservation of energy and momentum. In general, we know that an n degree of freedom mechanical system can be described by n its generalized coordinates $\{q_1, q_2, ..., q_n\}$*. An arbitrary conserved quantity of the system is also known as an* ***integral of motion****, and can be written as:*

$$\frac{d}{dt} I(t, q_1, q_1, ..., q_n, \dot{q}_1, \dot{q}_2, ...\dot{q}_n) = 0 \tag{4.1}$$

$$\Rightarrow \; I(t, q_1, q_1, ..., q_n, \dot{q}_1, \dot{q}_2, ...\dot{q}_n) = \text{constant} \tag{4.2}$$

We will begin our adventure into conserved quantities by first looking at a simple example.

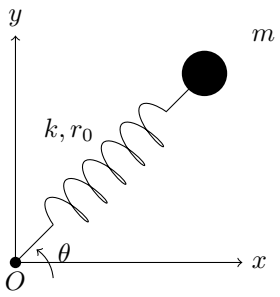

Figure 4.1: Spring-Mass System on a Plane

Example

Consider a spring-mass system, with a spring with spring constant k, rest length r_0 and the mass with mass m (ignoring gravity). This system has 2 degrees of freedom and as such, can be described by the 2 generalized coordinates $\{r, \theta\}$. We know from Newtonian mechanics that energy and angular momentum are conserved. These are written as:

1. Energy Conservation:

$$\begin{aligned} E &= T + V \\ &= \frac{1}{2}m\left(\dot{r}^2 + r^2\dot{\theta}^2\right) + \frac{1}{2}k(r - r_0)^2 = C_1 \end{aligned} \tag{4.3}$$

2. Angular Momentum Conservation:

$$\begin{aligned} & \dot{\vec{H}}_O = 0 \\ \Rightarrow\ & \vec{H}_O = \text{constant} \\ \Rightarrow\ & \vec{H}_O = \vec{r} \times m\vec{v} = r\vec{e}_r \times m(\dot{r}\vec{e}_r + r\dot{\theta}\vec{e}_\theta) \\ \Rightarrow\ & \vec{H}_O = mr^2\dot{\theta}\vec{e}_z = \text{constant} \\ \Rightarrow\ & mr^2\dot{\theta} = C_2 \end{aligned} \tag{4.4}$$

where C_1 and C_2 are constants. These 2 conservation laws in fact **complete** describe the motion of the system! Furthermore, the angular momentum conservation also elucidates a particular physical insight. That is, if the mass moves closer to the origin O, the angular velocity has to increase and vice versa. We will now look to how Lagrangian mechanics gives a natural segue to deriving these quantities and a generalization of this to all mechanical systems.

§4.1 Momentum Conservation

In this section, we will be looking at a conserved quantity that arises through *coordinate-translation invariance*. We start with a definition.

Definition 4.1.1. Cyclic Coordinates: *Generalized coordinates that do **not** explicitly occur in the Lagrangian of the system.*

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

$$\frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{q}_j} - \frac{\partial \mathcal{L}}{\partial q_j} = 0, \quad j \in \{1, ..., n\} \tag{4.5}$$

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 is:

$$\frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{q}_k} = 0 \tag{4.6}$$

because $\frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{q}_k} = 0$. So immediately we see that we have a conserved quantity here! We thus define this conserved quantity as the *generalized momentum*.

Definition 4.1.2. Generalized Momentum: *The generalized momentum of a system is defined as the partial derivative of the Lagrangian with respect to generalized velocity associated to a cycle coordinate q_k.*

$$p_k = \frac{\partial \mathcal{L}}{\partial \dot{q}_k} \tag{4.7}$$

§4.2 Hamiltonian Conservation

In this next section, we look at another conserved quantity which arises from *time-translation invariance*. Often times, this quantity turns out to be the total mechanical energy of a system, but is not strictly always the case. First consider an n degree of freedom system such that its Lagrangian is **not** an explicit function of time. That is to say:

$$\mathcal{L} = \mathcal{L}(q_1, ..., q_n, \dot{q}_1, ..., \dot{q}_n) \tag{4.8}$$

$$\Rightarrow \frac{\partial \mathcal{L}}{\partial t} = 0 \tag{4.9}$$

The *claim* is that from this, we get another integral of motion $\sum_j \frac{\partial \mathcal{L}}{\partial \dot{q}_j}\dot{q}_j - \mathcal{L} =$ constant. Our goal now is to prove/derive this. Keep in mind that $\frac{\partial \mathcal{L}}{\partial t} = 0$ does **not** imply $\mathcal{L}$ is an integral of motion since here, we are taking just the partial and not the total time derivative.

Claim

Given a system with a Lagrangian that is not an explicit function of time, we will have the following integral of motion:

$$\sum_j \frac{\partial \mathcal{L}}{\partial \dot{q}_j}\dot{q}_j - \mathcal{L} \tag{4.10}$$

Proof. Let us start by first taking the total time derivative of Lagrangian:

$$\frac{d}{dt}\mathcal{L} = \sum_{j=1}^{n} \frac{\partial \mathcal{L}}{\partial \dot{q}_j}\ddot{q}_j + \sum_{j=1}^{n} \frac{\partial \mathcal{L}}{\partial q_j}\dot{q}_j + \frac{\partial \mathcal{L}}{\partial t} \tag{4.11}$$

Also once again recall the Euler-Lagrange equation for conservative systems:

$$\begin{aligned} & \frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{q}_j} = \frac{\partial \mathcal{L}}{\partial q_j} \\ \Rightarrow \quad & \frac{d}{dt}\mathcal{L} = \sum_{j=1}^{n} \frac{\partial \mathcal{L}}{\partial \dot{q}_j}\ddot{q}_j + \sum_{j=1}^{n} \frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{q}_j}\dot{q}_j + \frac{\partial \mathcal{L}}{\partial t} \\ \Rightarrow \quad & \frac{d}{dt}\mathcal{L} = \frac{d}{dt}\left(\sum_{j=1}^{n} \frac{\partial \mathcal{L}}{\partial \dot{q}_j}\dot{q}_j\right) + \frac{\partial \mathcal{L}}{\partial t} \\ \Rightarrow \quad & \frac{d}{dt}\left(\sum_{j=1}^{n} \frac{\partial \mathcal{L}}{\partial \dot{q}_j}\dot{q}_j - \mathcal{L}\right) = -\frac{\partial \mathcal{L}}{\partial t} \end{aligned} \tag{4.12}$$

So we see that if our Lagrangian is indeed not an explicit function of time, we get that:

$$\frac{d}{dt}\left(\sum_{j=1}^{n} \frac{\partial \mathcal{L}}{\partial \dot{q}_j}\dot{q}_j - \mathcal{L}\right) = 0 \tag{4.13}$$

which proves our claim. □

However, this quantity seems kind of complicated and not very insightful for us. We will now work toward showing that this quantity is in fact (for the specific case where position is fully specified using only the generalized coordinates) total mechanical energy of our system. This integral of motion in the most general form is known as the *Hamiltonian*. Recall that $\mathcal{L} = T - V$, where V only depends on the generalized coordinates. That is to say:

$$\begin{aligned} & V = V(q_1, ..., q_n) \\ \Rightarrow \quad & \frac{\partial V}{\partial \dot{q}_j} = 0 \\ \Rightarrow \quad & \frac{\partial \mathcal{L}}{\partial \dot{q}_j} = \frac{\partial T}{\partial \dot{q}_j} \\ \Rightarrow \quad & \sum_{j=1}^{n} \frac{\partial T}{\partial \dot{q}_j}\dot{q}_j - T + V = \text{constant} \end{aligned} \tag{4.14}$$

Then for a system of particles where the vector coordinates are only functions of the generalized coordinates (i.e. $\vec{r}_j = \vec{r}_j(q_1, ..., q_n)$), we have:

$$\begin{aligned} \frac{\partial T}{\partial \dot{q}_j} &= \frac{\partial}{\partial \dot{q}_j}\sum_i \frac{m_i}{2}\vec{v}_i \cdot \vec{v}_i \\ &= \sum_i \frac{m_i}{2}\frac{\partial \vec{v}_i}{\partial \dot{q}_j} \cdot \vec{v}_i \\ &= \sum_i \frac{m_i}{2}\frac{\partial \vec{r}_i}{\partial q_j} \cdot \vec{v}_i \end{aligned} \tag{4.15}$$

Combining this with our previous result, we get:

$$\begin{aligned}\sum_{j=1}^{n}\frac{\partial T}{\partial \dot{q}_j}\dot{q}_j &= \sum_{j=1}^{n}\sum_{i} m_i \dot{q}_j \frac{\partial \vec{r}_i}{\partial q_j}\cdot\vec{v}_i \\ &= \sum_{i} m_i \vec{v}_i\cdot\vec{v}_i = 2T\end{aligned} \tag{4.16}$$

$$\begin{aligned}\Rightarrow \quad \sum_{j=1}^{n}\frac{\partial T}{\partial \dot{q}_j}\dot{q}_j - T + V &= 2T - T + V \\ &= T + V = \text{constant}\end{aligned} \tag{4.17}$$

where $T + V$ is exactly the total mechanical energy of the system and thus, energy is indeed conserved for systems with a Lagrangian not being an explicit function of time and positions being specified by generalized coordinates. Let us now consider an example to illustrate identifying these conserved quantities in a given mechanical system.

Example

Consider again the hinged, spring-mass system that we saw in 4.1. The first step is to write down the Lagrangian of this system. We have already previously identified the kinetic and potential energies in this system using the 2 generalized coordinates r and θ. As such, we have:

$$\mathcal{L} = T - V = \frac{1}{2}m(\dot{r}^2 + r^2\dot{\theta}^2) - \frac{1}{2}k(r - r_0)^2 \tag{4.18}$$

From here, we see that this Lagrangian is a function of only $r, \dot{r}, \dot{\theta}$ (i.e $\mathcal{L} = \mathcal{L}(r, \dot{r}, \dot{\theta})$). This means that by our derivations in this chapter, we have the 2 conserved quantities associated to θ and t which this Lagrangian is not an explicit function of. As such, this gives us that the following quantities are conserved:

$$\frac{\partial \mathcal{L}}{\partial t} = 0 \quad \Rightarrow \quad \frac{\partial \mathcal{L}}{\partial \dot{r}}\dot{r} + \frac{\partial \mathcal{L}}{\partial \dot{\theta}}\dot{\theta} - \mathcal{L} = C_1 \tag{4.19}$$

$$\frac{\partial \mathcal{L}}{\partial \theta} = 0 \quad \Rightarrow \quad \frac{\partial \mathcal{L}}{\partial \dot{\theta}} = C_2 \tag{4.20}$$

Working this out, we get:

$$\frac{1}{2}m(\dot{r}^2 + r^2\dot{\theta}^2) + \frac{1}{2}k(r - r_0)^2 = C_1 \tag{4.21}$$

$$mr^2\dot{\theta} = C_2 \tag{4.22}$$

Which are exactly the total mechanical energy (4.21) and angular momentum (4.22) of the system that we knew (from Newtonian mechanics) would be conserved as said at the beginning of the chapter.

To give a quick overview of what we've seen in this chapter, we have seen 2 conserved quantities arise from 2 separate conditions in the Lagrangian of a given mechanical system. The conditions and associated conserved quantities are summarized below:

1. **Conservation of Generalized Momentum**
   Condition: Lagrangian is not an explicit function of generalized coordinate q_k (i.e. q_k is a cyclic coordinate).

$$\begin{aligned} &\frac{\partial \mathcal{L}}{\partial q_k} = 0 \\ \Rightarrow \quad &\boxed{\frac{\partial \mathcal{L}}{\partial \dot{q}_k} = \text{constant}} \end{aligned} \tag{4.23}$$

2. **Conservation of the Hamiltonian**
   Condition: Lagrangian is not an explicit function of time t.

$$\begin{aligned} &\frac{\partial \mathcal{L}}{\partial t} = 0 \\ \Rightarrow \quad &\boxed{\sum_j \frac{\partial \mathcal{L}}{\partial \dot{q}_j} \dot{q}_j - \mathcal{L} = \text{constant}} \end{aligned} \tag{4.24}$$

Chapter 5

Oscillations

In this chapter, we will be looking at oscillators starting from the most ideal and simple scenario. This is most commonly known as simple harmonic motion. From there, we will then progressively work up to analysis for more complicated oscillatory systems. A good grasp of oscillator dynamics is essential for all scientists and engineers, with its uses proven invaluable time and time again.

§5.1 Single Mode Slit

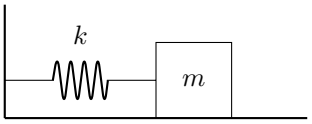

Figure 5.1: Spring-Mass System

Consider a simple spring-mass system, with parameters k (spring constant) and m (mass) as seen in figure 5.1. The equation of motion for this system is given as:

$$m\ddot{x} = -kx \tag{5.1}$$

for which we have seen before many times. We can solve for a general solution by taking the ansatz as:

$$x(t) = A\cos(\omega t) + B\sin(\omega t) \tag{5.2}$$

Plugging this back into the simple harmonic equation of motion above, we get:

$$\begin{aligned} \text{LHS} &= m\frac{d^2}{dt^2}\left(A\cos(\omega t) + B\sin(\omega t)\right) \\ &= -m\left(\omega^2 A\cos(\omega t) + \omega^2 B\sin(\omega t)\right) \\ &= -m\omega^2 x(t) \end{aligned} \tag{5.3}$$

Comparing this with the RHS of the equation, we get that:

$$\begin{aligned} m\omega^2 &= k \\ \Rightarrow\ \omega &= \sqrt{\frac{k}{m}} \end{aligned} \tag{5.4}$$

Now let's say that the initial conditions of our system as given as:

$$\begin{cases} x(t=0) = x_0 \\ \dot{x}(t=0) = \dot{x}_0 \end{cases} \tag{5.5}$$

With this, if we plug in $t = 0$ into our ansatz and compare coefficients with the given initial conditions, we will see that our general solution will take the form:

$$\boxed{x(t) = x_0 \cos(\omega t) + \frac{\dot{x}_0}{\omega} \sin(\omega t)} \tag{5.6}$$

This is a rather simple and elegant form for a general solution to simple harmonic systems. However, do not aim to memorize this but rather, keep in mind the ansatz we started with (linear combination of sines and cosines) as it will be essential for our analysis of oscillatory systems throughout this chapter.

§5.2 Damped Oscillators

Now consider a again a spring-mass system, but now with a damper added in parallel to the spring. A visualization of this is provided in figure 5.2 below.

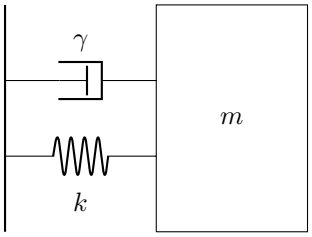

Figure 5.2: Spring-Damper-Mass System

The damping force is proportional to the velocity $\dot{x}$ of the mass with a proportionality constant $\gamma/2$. As such, the equation of motion for this system is:

$$\ddot{x} + 2\gamma\dot{x} + \frac{k}{m}x = 0 \tag{5.7}$$

To solve this differential equation, we shall use the following ansatz:

$$x(t) = Ce^{\lambda t} \tag{5.8}$$

Substituting this ansatz into the equation of motion, we get:

$$C(\lambda^2 + 2\gamma\lambda + \frac{k}{m})e^{\lambda t} = 0 \tag{5.9}$$

Recall that without the damper, we had that the characteristic angular frequency of a spring-mass system is given by $\sqrt{k/m}$. As such, we shall define:

$$\omega_0 = \sqrt{\frac{k}{m}} \tag{5.10}$$

With this, we solve the quadratic equation above for λ to get:

$$\boxed{\lambda_{1,2} = -\gamma \pm \sqrt{\gamma^2 - \omega_0^2}} \tag{5.11}$$

So we have 2 solutions to for λ and as such, the general solution to our second order differential equation (equation of motion) can be constructed by using both these quantities as follows:

$$x(t) = C_1 e^{\lambda_1 t} + C_2 e^{\lambda_2 t} \tag{5.12}$$

$$= C_1 e^{\left(-\gamma + \sqrt{\gamma^2 - \omega_0^2}\right)t} + C_2 e^{\left(-\gamma - \sqrt{\gamma^2 - \omega_0^2}\right)t} \tag{5.13}$$

where C_1 and C_2 can be found via knowing the initial conditions of our system. Let's dig a little deeper into this system. There are 3 scenarios we can look at here.

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

   Here, we see that our solution parameter becomes $\lambda_{1,2} = \pm i\omega_0$. With this, we simply retrieve the result previously attained for the case of no damping. This case is known as **no damping**.

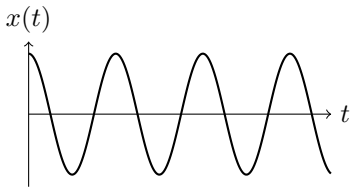

Figure 5.3: No Damping

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

   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**.

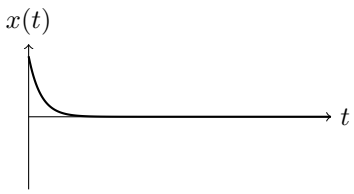

Figure 5.4: Over Damping

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

   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 $\omega = \sqrt{\omega_0^2 - \gamma^2}$.

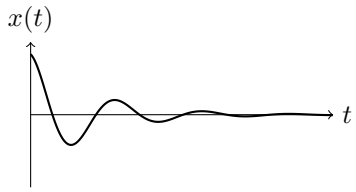

Figure 5.5: Under Damping

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

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

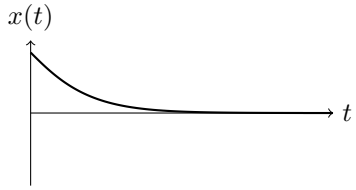

Figure 5.6: Critical Damping

   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:

$$x(t) = e^{-\gamma t}(A + Bt) \tag{5.14}$$

§5.3 Forced Oscillations

We now continue to augment our system for further analysis by adding an external time-dependent force on the mass.

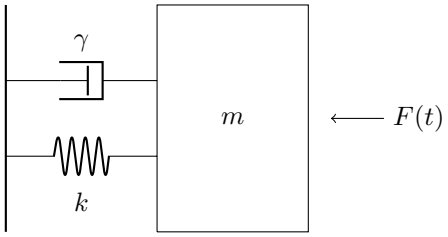

Figure 5.7: Spring-Damper-Mass System

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

$$F(t) = F_0 \sin(\omega t) \tag{5.15}$$

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:

$$\ddot{x} + 2\gamma\dot{x} + \frac{k}{m}x = \frac{F_0}{m}\sin(\omega t) \tag{5.16}$$

where we defined γ 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:

$$\begin{aligned} x(t) &= x_h(t) + x_p(t) \\ &= \left[C_1 e^{\left(-\gamma+\sqrt{\gamma^2-\omega_0^2}\right)t} + C_2 e^{\left(-\gamma-\sqrt{\gamma^2-\omega_0^2}\right)t}\right] + [A\sin(\omega t) + B\cos(\omega t)] \end{aligned} \tag{5.17}$$

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

$$\ddot{x}_p(t) + 2\gamma\dot{x}_p(t) + \frac{k}{m}x_p(t) = \frac{F_0}{m}\sin(\omega t) \tag{5.18}$$

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 $x_p(t)$:

$$\begin{aligned} & (\omega_0^2 - \omega^2)\left[A\sin(\omega t) + B\cos(\omega t)\right] + 2\gamma\omega\left[A\cos(\omega t) - B\sin(\omega t)\right] = \frac{F_0}{m}\sin(\omega t) \\ \Rightarrow\ & \left[(\omega_0^2 - \omega^2)A - 2\gamma\omega B\right]\sin(\omega t) + \left[(\omega_0^2 - \omega^2)B + 2\gamma\omega A\right]\cos(\omega t) = \frac{F_0}{m}\sin(\omega t) \end{aligned} \tag{5.19}$$

$$\begin{aligned} \Rightarrow\ & (\omega_0^2 - \omega^2)A - 2\gamma\omega B = \frac{F_0}{m} \\ & (\omega_0^2 - \omega^2)B + 2\gamma\omega A = 0 \end{aligned} \tag{5.20}$$

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

$$\frac{1}{T}\int_0^T \sin(\omega t)\cos(\omega t)dt = 0 \tag{5.21}$$

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:

$$\begin{bmatrix} \omega_0^2 - \omega^2 & -2\gamma\omega \\ 2\gamma\omega & \omega_0^2 - \omega^2 \end{bmatrix} \begin{bmatrix} A \\ B \end{bmatrix} = \begin{bmatrix} \frac{F_0}{m} \\ 0 \end{bmatrix} \tag{5.22}$$

$$\Rightarrow \quad \begin{aligned} A &= \frac{\frac{F_0}{m}(\omega_0^2 - \omega^2)}{(\omega_0^2 - \omega^2)^2 + (2\gamma\omega)^2} \\ B &= \frac{-2\gamma\omega\frac{F_0}{m}}{(\omega_0^2 - \omega^2)^2 + (2\gamma\omega)^2} \end{aligned} \tag{5.23}$$

$$\Rightarrow \; x_p(t) = \frac{\left(\frac{F_0}{m}\right)}{(\omega_0^2 - \omega^2)^2 + (2\gamma\omega)^2} \left[(\omega_0^2 - \omega^2)\sin(\omega t) - 2\gamma\omega\cos(\omega t)\right] \tag{5.24}$$

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 some phase given by $\tan^{-1}(-B/A)$. Utilizing this, we get that our particular solution becomes:

$$x_p(t) = \frac{\left(\frac{F_0}{m}\right)}{(\omega_0^2 - \omega^2)^2 + (2\gamma\omega)^2} \sin\left[\omega t + \tan^{-1}\left(\frac{2\gamma\omega}{\omega_0^2 - \omega^2}\right)\right] \tag{5.25}$$

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

$$\begin{aligned} x(t) = &\left[C_1 e^{\left(-\gamma + \sqrt{\gamma^2 - \omega_0^2}\right)t} + C_2 e^{\left(-\gamma - \sqrt{\gamma^2 - \omega_0^2}\right)t}\right] \\ &+ \frac{\left(\frac{F_0}{m}\right)}{(\omega_0^2 - \omega^2)^2 + (2\gamma\omega)^2} \sin\left[\omega t + \tan^{-1}\left(\frac{2\gamma\omega}{\omega_0^2 - \omega^2}\right)\right] \end{aligned} \tag{5.26}$$

As further analysis, we can generate a plot of the amplitude of $x_p(t)$ against ω (as shown in figure 5.8 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.

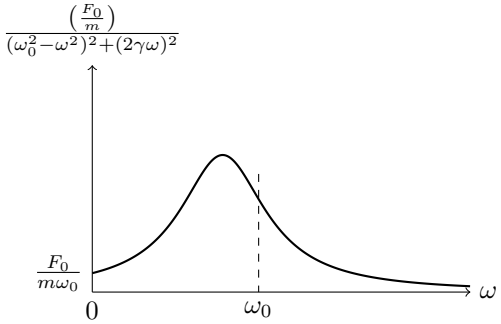

Figure 5.8: Oscillation Amplitude vs Angular Frequency Plot

There are several things we can already see from just eyeballing the plot above.

- We see that if $\omega >> \omega_0$, we get that $\frac{\left(\frac{F_0}{m}\right)}{(\omega_0^2-\omega^2)^2+(2\gamma\omega)^2} \to 0$.
- We see that the maximum amplitude does **not** occur at ω_0, although close by. The maximum amplitude only coincides with ω_0 when $\gamma = 0$.
- When there is no driving frequency ($\omega_0 = 0$), the driving amplitude does not vanish but tends to $F_0/(m\omega_0)$ instead.

§5.4 Multi-Mode Oscillators

Now, we will take a step back again removing damping and external forces on our oscillator system. However, we will instead be looking at oscillators of multiple modes (a system of several 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:

$$T = \sum_j \frac{1}{2} m_j \dot{q}_j^2 \tag{5.27}$$

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

$$V \approx \sum_{ij} \frac{1}{2} k_{ij} q_i q_j \tag{5.28}$$

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:

$$\ddot{q}_i + \frac{k_{ii}}{m_i} q_i + \sum_{j \neq i} \frac{k_{ij}}{2m_i} q_j = 0 \tag{5.29}$$

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:

$$\omega^2 q_i = \frac{k_{ii}}{m_i} q_i + \sum_{j \neq i} \frac{k_{ij}}{2m_i} q_j \tag{5.30}$$

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:

$$\begin{bmatrix} \ddots & & & & \\ & & \frac{k_{i-1,i+1}}{2m_i} & & \\ & \frac{k_{i,i}}{m_i} & & & \\ & \frac{k_{i+1,i-1}}{2m_i} & & & \\ & & & & \ddots \end{bmatrix} \begin{bmatrix} \vdots \\ q_{i-1} \\ q_i \\ q_{i+1} \\ \vdots \end{bmatrix} = \omega^2 \begin{bmatrix} \vdots \\ q_{i-1} \\ q_i \\ q_{i+1} \\ \vdots \end{bmatrix} \tag{5.31}$$

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:

$$(M^{-1}K)\vec{q} = \omega^2\vec{q} \tag{5.32}$$

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 $\vec{q}$). 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:

   $$q(t) = Ae^{i\omega t} + Be^{-i\omega t} \tag{5.33}$$

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:

   $$q(t) = Ae^{\omega t} + Be^{-\omega t} \tag{5.34}$$

   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:

   $$q(t) = \frac{\varepsilon}{2}(e^{\omega t} + e^{-\omega t}) = \varepsilon \sinh(\omega t) \tag{5.35}$$

   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.

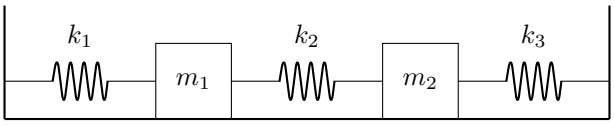

Figure 5.9: Multi-Mode System of Springs and Masses

Example:

Consider the following coupled system of springs and masses as seen in figure 5.9. With this, we can write the equations of motion as:

$$m_1\ddot{x}_1 = -k_1x_1 + k_2(x_2 - x_1) \tag{5.36}$$
$$m_2\ddot{x}_2 = -k_2(x_2 - x_1) + k_3x_2 \tag{5.37}$$

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:

$$\begin{bmatrix} m_1 & 0 \\ 0 & m_2 \end{bmatrix} \begin{bmatrix} \ddot{x}_1 \\ \ddot{x}_2 \end{bmatrix} = \begin{bmatrix} -k_1 - k_2 & k_2 \\ k_2 & -k_2 - k_3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \tag{5.38}$$

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

$$\vec{x} = e^{i\omega t}\vec{x}_0 \tag{5.39}$$
$$M = \begin{bmatrix} m_1 & 0 \\ 0 & m_2 \end{bmatrix}, \quad K = \begin{bmatrix} k_1 + k_2 & -k_2 \\ -k_2 & k_2 + k_3 \end{bmatrix} \tag{5.40}$$

where $\vec{x}_0$ 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:

$$\begin{aligned} & \omega^2 M\vec{x}_0 = K\vec{x}_0 \\ \Rightarrow \quad & M^{-1}K\vec{x}_0 = \omega^2\vec{x}_0 \end{aligned} \tag{5.41}$$

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:

$$\begin{aligned} & \det\left(M^{-1}K - \mathbb{I}\omega^2\right) = 0 \\ \Rightarrow \quad & \det\left\{\frac{k}{m}\begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix} - \omega^2\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\right\} = 0 \\ \Rightarrow \quad & (2 - \frac{m\omega^2}{k})^2 - 1 = 0 \\ \Rightarrow \quad & \frac{m\omega^2}{k} = 1 \text{ or } 3 \end{aligned} \tag{5.42}$$

$$\Rightarrow \quad \omega = \sqrt{\frac{3k}{m}} \text{ or } \sqrt{\frac{k}{m}} \tag{5.43}$$

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

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

$$\begin{aligned} & \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} - \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \vec{0} \\ \Rightarrow \; & x_1 = x_2 \\ \Rightarrow \; & \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \end{aligned} \tag{5.44}$$

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

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

$$\begin{aligned} & \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} - 3 \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \vec{0} \\ \Rightarrow \; & x_1 = -x_2 \\ \Rightarrow \; & \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \end{bmatrix} \end{aligned} \tag{5.45}$$

   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:

$$\boxed{\vec{x}(t) = \sum_j A_j e^{i\omega_j t} \vec{\eta}_j} \tag{5.46}$$

where A_j are real coefficients and $\vec{\eta}_j$ are the eigen-modes.

§5.5 Beat Phenomena

In oscillatory systems, a unique phenomena can emerge in which the periodic variations in vibrational amplitudes occur at a rate that is the difference of two characteristic frequencies $(\omega_1 - \omega_2)$ of the system under certain 'tuned' conditions. This is known as *beat phenomenon*, and arises from the interference between two oscillations of **slightly** different frequencies. Mathematically, this simply results from the properties of trigonometric functions, but reveals rather insightful physics. We will be exploring how this arises in the dynamics of the *Wilberforce pendulum*.

§5.5.1 Wilberforce Pendulum

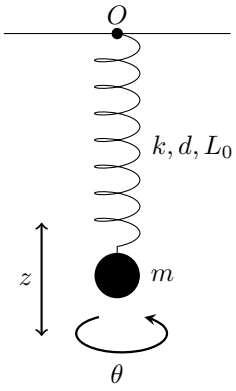

Figure 5.10: Wilberforce Pendulum

Consider a mass hanging on a spring that is free to oscillate in both the standard longitudinal mode and the torsional mode. When the mass is lifted above its equilibrium point and released from rest, it oscillates up and down along a vertical line, but due to the coupling between this longitudinal and the torsional motion, rotation of the mass can also be observed. This constitutes the Wilberforce pendulum and is illustrated in figure 5.10, where k is the longitudinal stiffness, d is the torsional stiffness and L_0 is the rest length of the spring. For simplicity, we will consider a linear coupling between the torsional and longitudinal modes parameterized by the constant ε. As such, our energy functions are given by:

$$T = \frac{1}{2}m\dot{z}^2 + \frac{1}{2}I\dot{\theta}^2 \tag{5.47}$$

$$\begin{aligned} V &= \frac{1}{2}k(z - L_0)^2 + \frac{1}{2}d\theta^2 + \frac{\varepsilon}{2}\theta z - mgz \\ &= \frac{1}{2}kz^2 + \frac{1}{2}d\theta^2 + z\left(\frac{\varepsilon}{2}\theta - kL_0 - mg\right) + \frac{1}{2}kL_0^2 \end{aligned} \tag{5.48}$$

The potential energy function here is slightly inconvenient since it would result in an affine system as opposed to a purely linear one. As such, we can do an appropriate coordinate transformation (change of reference frame) such that we simplify our computations. To do so, let us define our new coordinates as:

$$\tilde{z} = z + z_0 \tag{5.49}$$

$$\tilde{\theta} = \theta + \theta_0 \tag{5.50}$$

where $\tilde{z}$ and $\tilde{\theta}$ are the coordinates in the new frame and z_0 and θ_0 are the associated 'boosts' (transformation factors). The goal is to have z_0 and θ_0 such that our potential energy function reduces to:

$$\tilde{V}(\tilde{\theta}, \tilde{z}) = \frac{1}{2}k\tilde{z}^2 + \frac{1}{2}d\tilde{\theta}^2 + \frac{\varepsilon}{2}\tilde{\theta}\tilde{z} = V(\theta, z) \tag{5.51}$$

As such, we substitute in our newly proposed generalized coordinates into $\tilde{V}$ to get:

$$\begin{aligned}\tilde{V}(\theta,z) &= \frac{1}{2}kz^2 + kzz_0 + \frac{1}{2}kz_0^2 \\ &+ \frac{1}{2}d\theta^2 + d\theta\theta_0 + \frac{1}{2}d\theta_0^2 \\ &+ \frac{\varepsilon}{2}\theta z + \frac{\varepsilon}{2}(\theta z_0 + \theta_0 z) + \frac{\varepsilon}{2}\theta_0 z_0 \\ &= \frac{1}{2}kz^2 + \frac{1}{2}d\theta^2 \\ &+ z\left(\frac{\varepsilon}{2}\theta + \frac{\varepsilon}{2}\theta_0 + kz_0\right) + \theta\left(\frac{\varepsilon}{2}z_0 + d\theta_0\right) \\ &+ \left(\frac{1}{2}zL_0^2 + \frac{1}{2}d\theta_0^2 + \frac{\varepsilon}{2}\theta_0 z_0\right)\end{aligned} \tag{5.52}$$

From here, we compare $\tilde{V}(\theta,z)$ and $V(\theta,z)$. Let's write them side by side for clarity.

$$\begin{aligned}\tilde{V}(\theta,z) &= \frac{1}{2}kz^2 + \frac{1}{2}d\theta^2 \\ &+ z\left(\frac{\varepsilon}{2}\theta + \frac{\varepsilon}{2}\theta_0 + kz_0\right) + \theta\left(\frac{\varepsilon}{2}z_0 + d\theta_0\right) \\ &+ \left(\frac{1}{2}zL_0^2 + \frac{1}{2}d\theta_0^2 + \frac{\varepsilon}{2}\theta_0 z_0\right)\end{aligned} \qquad \begin{aligned}V(\theta,z) &= \frac{1}{2}kz^2 + \frac{1}{2}d\theta^2 \\ &+ z\left(\frac{\varepsilon}{2}\theta - kL_0 - mg\right) \\ &+ \frac{1}{2}kL_0^2\end{aligned} \tag{5.53}$$

From this, we simply compare the coefficients of z and θ to solve for z_0 and θ_0 (ignoring the constant term since it drops out of the Lagrangian):

$$\theta_0 = \frac{L_0 + \frac{mg}{k}}{\frac{2d}{\varepsilon} - \frac{\varepsilon}{2k}}, \quad z_0 = \frac{2d\left(L_0 + \frac{mg}{k}\right)}{\varepsilon\left(\frac{\varepsilon}{2k} - \frac{2d}{\varepsilon}\right)} \tag{5.54}$$

Great, now we have that our Lagrangian in this new coordinate system is effectively given by:

$$\begin{aligned}\mathcal{L} &= T(\dot{\tilde{\theta}},\dot{\tilde{z}}) - \tilde{V}(\tilde{\theta},\tilde{z}) \\ &= \left[\frac{1}{2}m\dot{\tilde{z}}^2 + \frac{1}{2}I\dot{\tilde{\theta}}^2\right] - \left[\frac{1}{2}k\tilde{z}^2 + \frac{1}{2}d\tilde{\theta}^2 + \frac{\varepsilon}{2}\tilde{\theta}\tilde{z}\right]\end{aligned} \tag{5.55}$$

To reduce notation, we will drop the tilde on our generalized coordinates which simply gives:

$$\mathcal{L} = \left[\frac{1}{2}m\dot{z}^2 + \frac{1}{2}I\dot{\theta}^2\right] - \left[\frac{1}{2}kz^2 + \frac{1}{2}d\theta^2 + \frac{\varepsilon}{2}\theta z\right] \tag{5.56}$$

From here, we derive our conjugate forces and momenta using the standard methodology and plug these into the Euler-Lagrange equation to get the following equations of motion:

$$m\ddot{z} + kz + \frac{1}{2}\varepsilon\theta = 0 \tag{5.57}$$

$$I\ddot{\theta} + d\theta + \frac{1}{2}\varepsilon z = 0 \tag{5.58}$$

$$\Rightarrow \quad \begin{bmatrix} m & 0 \\ 0 & I \end{bmatrix}\begin{bmatrix}\ddot{z} \\ \ddot{\theta}\end{bmatrix} = -\begin{bmatrix} k & \frac{\varepsilon}{2} \\ \frac{\varepsilon}{2} & d \end{bmatrix}\begin{bmatrix} z \\ \theta \end{bmatrix} \tag{5.59}$$

We then plug in the multi-mode oscillator ansatz to get:

$$\begin{bmatrix} \frac{k}{m} & \frac{\varepsilon}{2m} \\ \frac{\varepsilon}{2I} & \frac{d}{I} \end{bmatrix} \begin{bmatrix} z \\ \theta \end{bmatrix} = \omega^2 \begin{bmatrix} z \\ \theta \end{bmatrix} \tag{5.60}$$

The solutions for this eigenvalue problem are then the normal modes and eigen-frequencies of our double-mode oscillator system, given as:

$$\omega_{\pm}^2 = \frac{1}{2}\left[\left(\frac{k}{m}+\frac{d}{I}\right) \pm \sqrt{\left(\frac{k}{m}+\frac{d}{I}\right)^2 - 4\left(\frac{dk}{mI}+\frac{\varepsilon^2}{4mI}\right)}\right] \tag{5.61}$$

To further simplify our analysis, let's say we tune the parameters of our system in such a way that $\frac{k}{m} = \frac{d}{I}$. This reduces our eigen-frequencies to:

$$\omega_{\pm}^2 = \frac{k}{m} \pm \frac{\varepsilon}{2\sqrt{mI}} \tag{5.62}$$

$$\Rightarrow \ \omega_{\pm} \approx \sqrt{\frac{k}{m}}\left(1 \pm \frac{m\varepsilon}{4k\sqrt{mI}}\right) \tag{5.63}$$

where we have assumed that ε is small and taken the Taylor expansion. These have the associated eigen-modes:

$$\vec{\eta}_+ = \begin{bmatrix} 1 \\ \sqrt{\frac{m}{I}} \end{bmatrix}, \quad \vec{\eta}_- = \begin{bmatrix} 1 \\ -\sqrt{\frac{m}{I}} \end{bmatrix} \tag{5.64}$$

The general solution of this system is thus given by:

$$\begin{bmatrix} z(t) \\ \theta(t) \end{bmatrix} = C_+ \begin{bmatrix} 1 \\ \sqrt{\frac{m}{I}} \end{bmatrix} \cos(\omega_+ t + \phi_+) + C_- \begin{bmatrix} 1 \\ -\sqrt{\frac{m}{I}} \end{bmatrix} \cos(\omega_- t + \phi_-) \tag{5.65}$$

where $\phi_\pm$ are the phase shifts determined by the initial conditions of the system. Let's now say that we set-up the system with the following initial conditions:

$$\begin{bmatrix} z(0) \\ \theta(0) \end{bmatrix} = \begin{bmatrix} A \\ 0 \end{bmatrix}, \quad \begin{bmatrix} \dot{z}(0) \\ \dot{\theta}(0) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \tag{5.66}$$

which corresponds to initially stretching the spring by a length A (with no torsional rotation), then letting it go to allow free evolution of the undamped system. With these initial conditions, we solve for $\phi_\pm$ and $C_\pm$ to get:

$$C_+ = C_- = \frac{A}{2}, \quad \phi_+ = \phi_- = 0 \tag{5.67}$$

Rendering our dynamical solution as:

$$\begin{bmatrix} z(t) \\ \theta(t) \end{bmatrix} = \frac{A}{2}\begin{bmatrix} 1 \\ \sqrt{\frac{m}{I}} \end{bmatrix} \cos(\omega_+ t) + \frac{A}{2}\begin{bmatrix} 1 \\ -\sqrt{\frac{m}{I}} \end{bmatrix} \cos(\omega_- t) \tag{5.68}$$

$$\Rightarrow \begin{aligned} z(t) &= \frac{A}{2}\left[\cos(\omega_+ t) + \cos(\omega_- t)\right] \\ \theta(t) &= \frac{A}{2}\sqrt{\frac{m}{I}}\left[\cos(\omega_+ t) - \cos(\omega_- t)\right] \end{aligned} \tag{5.69}$$

Here is where the magic happens. First, recall the following trigonometric identities:

$$\begin{aligned}\cos(A)+\cos(B) &= 2\cos\left(\frac{A+B}{2}\right)\cos\left(\frac{A-B}{2}\right)\\ \cos(A)-\cos(B) &= 2\sin\left(\frac{A+B}{2}\right)\sin\left(\frac{A-B}{2}\right)\end{aligned} \tag{5.70}$$

With this, this gives us that our dynamical solutions can also be written as:

$$z(t) = A\cos\left[(\omega_+-\omega_-)t\right]\cdot\cos\left[(\omega_++\omega_-)t\right] = A\cos\left(t\sqrt{\frac{k}{m}}\right)\cdot\cos\left(t\frac{\varepsilon}{4\sqrt{kI}}\right) \tag{5.71}$$

$$\theta(t) = A\sqrt{\frac{m}{I}}\sin\left[(\omega_+-\omega_-)t\right]\cdot\sin\left[(\omega_++\omega_-)t\right] = A\sqrt{\frac{m}{I}}\sin\left(t\sqrt{\frac{k}{m}}\right)\cdot\sin\left(t\frac{\varepsilon}{4\sqrt{kI}}\right) \tag{5.72}$$

The resulting solutions will thus take on the forms as shown in figures 5.11 and 5.12.

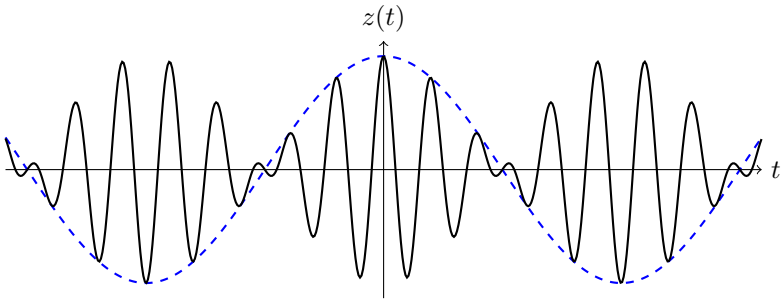

Figure 5.11: $z(t)$ vs t Plot

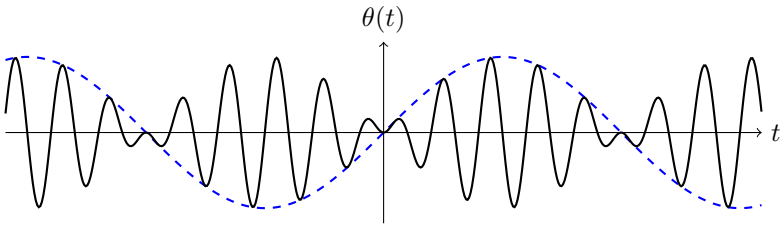

Figure 5.12: $\theta(t)$ vs t Plot

which exactly displays the emergence of beat phenomena we alluded to at the start of this section! Beat phenomena is seen not just in mechanical systems, but any physical system that consists of wave solutions. In fact, beat phenomena is what allows us to transmit information in the form of radiowaves through the air! It is also common place in the study of acoustics, where beat phenomenon causes fluctuations in the volume of sounds from some source.

Chapter 6

Kepler's Laws

From here on in the notes, advanced topics will be covered that are not usually touched on in the standard edition of 30.104 Dynamics. Naturally, these will require understanding of the material from the previous chapters as we will be studying more sophisticated systems and concepts. The first of these will be on Kepler's laws of planetary motion. These were formulated by Johannes Kepler between 1609 and 1619 while he was studying the motion of planets around the Sun.

§6.1 Statement of Kepler's Laws

Kepler's 3 laws that model the motion of Earth around the Sun are succinctly stated 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 chapter, 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.

§6.1.1 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. Following Newton's second law, the forces on these masses are given by:

$$M\ddot{\vec{r}}_M = \frac{GMm}{r^2}\vec{r}_r, \quad m\ddot{\vec{r}}_m = -\frac{GMm}{r^2}\vec{r}_r \tag{6.1}$$

$$\Rightarrow \quad \ddot{\vec{r}} = -\frac{G\mu}{r^2}\vec{e}_r \tag{6.2}$$

Where we define $\vec{r} = \vec{r}_m + \vec{r}_M$ and also $\mu = m + M$. Since we take that there is no net external force acting on this 2 mass system, we have:

$$\frac{d^2}{dt^2}(M\vec{r}_M + m\vec{r}_m) = \vec{0} \tag{6.3}$$

From the result of Newton's second law, we see that there is only a force along the $\vec{e}_r$ direction. But we know from our earlier study of dynamics that this is not true in general ($\ddot{\vec{r}} = (\ddot{r} - r\dot{\theta}^2)\vec{e}_r + (r\ddot{\theta} + 2\dot{r}\dot{\theta})\vec{e}_\theta$). As such, we get the term associated to the $\vec{e}_\theta$ direction must vanish:

$$\begin{aligned} &\Rightarrow\ r\ddot{\theta} + 2\dot{r}\dot{\theta} = 0 \\ &\Rightarrow\ \frac{1}{2}\frac{d}{dt}r^2\dot{\theta} = 0 \\ &\Rightarrow\ r^2\dot{\theta} = \text{constant} \\ &\Rightarrow\ \boxed{\frac{r^2}{2}\frac{d\theta}{dt} = \frac{dA}{dt} = \text{constant}} \end{aligned} \tag{6.4}$$

where we utilized the fact that $dA = r \cdot rd\theta/2$. From this, we see that the area of the orbit is indeed time-invariant, proving Kepler's second law.

§6.1.2 Kepler's First Law

The orbits of planets around the Sun are elliptical.

Taking a step back, we now proceed to derive Kepler's first law using the results obtained from the second. For convenience, we will be calling the conserved quantity earlier derived as $\frac{r^2}{2}\frac{d\theta}{dt} = \frac{dA}{dt} = \frac{h}{2}$. Utilizing the result obtained from center of mass frame, the $\vec{e}_r$ component gives us:

$$\ddot{r} - r\dot{\theta}^2 = -\frac{G\mu}{r^2} \tag{6.5}$$

From the earlier derivation, we have $r^2\dot{\theta} = h$ and hence, $r\dot{\theta}^2 = \frac{h^2}{r^3}$. Making use of a simple identity, we get:

$$\dot{r} = -r^2\frac{d}{dt}\frac{1}{r} = \frac{h}{\dot{\theta}}\frac{d}{dt}\frac{1}{r} = -h\frac{d}{d\theta}\frac{1}{r} \tag{6.6}$$

Taking the another derivative of this, we get:

$$\begin{aligned} \ddot{r} &= \frac{d}{dt}\left(-h\frac{d}{d\theta}\frac{1}{r}\right) \\ &= -\frac{h^2}{r^2}\frac{d^2}{d\theta^2}\frac{1}{r} \\ \Rightarrow\ -\frac{G\mu}{r^2} &= -\frac{h^2}{r^2}\frac{d^2}{d\theta^2}\frac{1}{r} - \frac{h^2}{r^3} \\ \Rightarrow\ \frac{G\mu}{h^2} &= \frac{d^2}{d\theta^2}\frac{1}{r} + \frac{1}{r} \end{aligned} \tag{6.7}$$

Solving this differential equation, we get:

$$\begin{aligned} \frac{1}{r} &= (e\cos\theta + 1)\frac{G\mu}{h^2} \\ \Rightarrow \quad &\boxed{r = \left(\frac{h^2}{G\mu}\right)\left(\frac{1}{1+e\cos\theta}\right)} \end{aligned} \tag{6.8}$$

For which e is known as the *eccentricity* and determines the orbit of the masses. 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.

§6.1.3 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 reasserting a result we have already found, that is $\frac{dA}{dt} = \frac{h}{2}$. Since we have an elliptical orbit, we have:

$$\pi ab = \frac{h}{2}T \tag{6.9}$$

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:

$$2a = r_1 + r_2 \tag{6.10}$$

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

$$\begin{aligned} & r(0) + r(\pi) = \frac{h^2}{G\mu}\frac{1}{1+e} + \frac{h^2}{G\mu}\frac{1}{1-e} = 2a \\ \Rightarrow \quad & a = \frac{h^2}{G\mu}\frac{1}{1-e^2} \\ \Rightarrow \quad & b = a\sqrt{1-e^2} \\ \Rightarrow \quad & \pi a(a\sqrt{1-e^2}) = \frac{T}{2}\sqrt{G\mu a(1-e^2)} \\ \Rightarrow \quad & \frac{T\sqrt{G\mu a}}{2} = \pi a \\ \Rightarrow \quad & \boxed{T^2 \propto a^3} \end{aligned} \tag{6.11}$$

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

§6.2 Lagrange Points in 3 Body Systems

Continuing on in the analysis of planetary motion, We will now look at a simplified model of the 3 body problem. The simplification we are adopting is that we are in the limit where one of the masses is negligible compared to the mass of the other 2 bodies. As such, we define our system by the 2 bodies of larger mass being M_1 and M_2, and the third body of much smaller mass being m. The assumption here is that the smaller mass does not affect the dynamics of the larger mass (but **not** vice versa). As such, we have:

$$\ddot{X}_1 = -\frac{GM_2(X_1 - X_2)}{\left((X_1 - X_2)^2 + (Y_1 - Y_2)^2\right)^{3/2}} \tag{6.12}$$

$$\ddot{X}_2 = -\frac{GM_1(X_2 - X_1)}{\left((X_1 - X_2)^2 + (Y_1 - Y_2)^2\right)^{3/2}} \tag{6.13}$$

$$\ddot{x}_3 = -\frac{GM_1(x_3 - X_1)}{\left((x_3 - X_1)^2 + (y_3 - Y_1)^2\right)^{3/2}} - \frac{GM_2(x_3 - X_2)}{\left((x_3 - X_2)^2 + (y_3 - Y_2)^2\right)^{3/2}} \tag{6.14}$$

where $\{X_j, Y_j\}$ is the position associated to M_j with respect to some global reference frame, and similar for $\{x_3, y_3\}$ to m. A further assumption we shall assert for simplicity is that the 2 larger masses undergo **circular** orbits about their collective center of mass. As such, $\sqrt{(X_1 - X_2)^2 + (Y_1 - Y_2)^2} = R$ which gives us:

$$(\ddot{X}_1 - \ddot{X}_2) = -\frac{G(M_1 + M_2)}{R^3}(X_1 - X_2) \tag{6.15}$$

$$\Rightarrow \ \Omega = \sqrt{\frac{G(M_1 + M_2)}{R^3}} \tag{6.16}$$

Now, let's consider an axis that rotates along with the 2 larger bodies, where the x-axis is along the line that joins the both of them with the origin at their center of mass (center of mass frame). we now study the dynamics of the smaller mass in this frame. We employ equation 1.17, for which we have the orthogonal unit vectors being $\vec{e}_x$ and $\vec{e}_y$ instead of $\vec{e}_r$ and $\vec{e}_\theta$. We thus have:

$$\frac{d^2x}{dt^2} - \Omega^2 x - 2\Omega\frac{dy}{dt} = -\frac{GM_1(x_3 - X_1)}{\left((x_3 - X_1)^2 + (y_3 - Y_1)^2\right)^{3/2}} - \frac{GM_2(x_3 - X_2)}{\left((x_3 - X_2)^2 + (y_3 - Y_2)^2\right)^{3/2}} \tag{6.17}$$

$$\frac{d^2y}{dt^2} - \Omega^2 y - 2\Omega\frac{dx}{dt} = -\frac{GM_1(y_3 - Y_1)}{\left((x_3 - X_1)^2 + (y_3 - Y_1)^2\right)^{3/2}} - \frac{GM_2(y_3 - Y_2)}{\left((x_3 - X_2)^2 + (y_3 - Y_2)^2\right)^{3/2}} \tag{6.18}$$

for which the $\Omega^2 x_j$ terms constitute centripetal acceleration, and the $2\Omega\dot{x}_j$ terms constitute Coriolis acceleration ($x_j \in \{x, y\}$). As a means to better understand the system, we look for *stable points* (solutions to $\boldsymbol{\nabla} V(x) = \vec{0}$). To do this, explore treatments to the *pseudopotential* (a.k.a *effective potential*). We first separate our potential into 2 parts:

$$\begin{aligned} V &= V_g - V_\Omega \\ &= \left[-\frac{GM_1}{\sqrt{(x + r_1)^2 + y^2}} - \frac{GM_2}{\sqrt{(x - r_2)^2 + y^2}}\right] - \left[\frac{\Omega^2}{2}(x^2 + y^2)\right] \end{aligned} \tag{6.19}$$

where V_g is the gravitational contribution to the potential and V_Ω the centrifugal. A visualization of this effective potential can be seen below in figure 6.1.

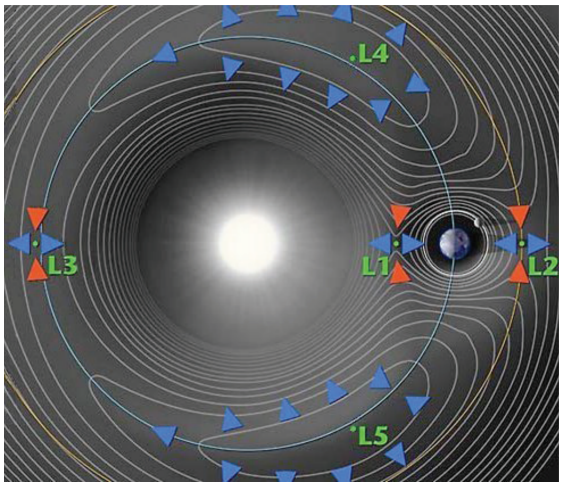

Figure 6.1: Effective Potential

As seen in the figure, there are certain points $(L_1, L_2, L_3, L_4, L_5)$ in this potential that due to the combined gravitational forces of 2 large bodies balancing with the centrifugal forces, this creates points of equilibrium. These equilibria are known as *Lagrange points*.

Chapter 7

Ideal Material Analysis

We have covered some rudimentary wave mechanics in chapter 5 in the form of discrete coupled oscillators. What we are about to do here is in essence generalize this to a continuum limit. We will see how this infinite series of coupled-differential harmonic oscillators can be used to nicely model some elastic bulk materials. After this, we will take a detour to inelastic but flexible materials, for which equilibrium configurations will be explored. Utilization of differential analysis and elementary knowledge of bulk material mechanics will be key in this chapter.

§7.1 Elastic Beams

Consider a beam of elastic material with total length L oriented along the x-axis and some cross-sectional area A. We also assert that the material has constant elasticity throughout its volume with a Young's modulus E. Now consider a differential segment of the beam of length dx as shown in figure 7.1 below.

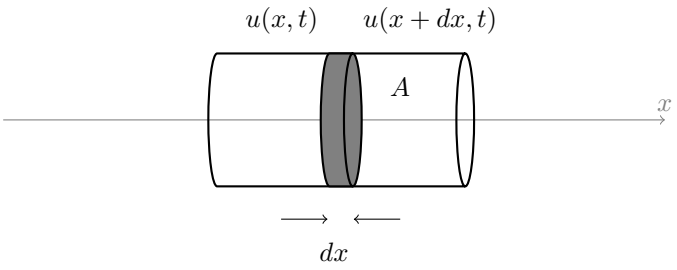

Figure 7.1: Differential Element of Elastic Beam

Above, we have taken $u(x)$ to be the x-displacement of the material at any point x along its length. As such, we have that:

$$\sigma = E\frac{\partial u(x,t)}{\partial x} \tag{7.1}$$

where σ is the applied axial stress on the material and $\frac{\partial u(x)}{\partial x}$ the strain. From here, we use Newton's second law on the differential segment, which gives us:

$$\begin{aligned} m\ddot{u} &= \sum F = A\sum \sigma \\ \Rightarrow\ \rho A dx \frac{\partial^2 u(x,t)}{\partial t^2} &= A\left(\sigma(x+dx) - \sigma(x)\right) \end{aligned} \tag{7.2}$$

where ρ is the density of the material. The right hand side of the expression above can be further written as:

$$\begin{aligned} A\left(\sigma(x+dx) - \sigma(x)\right) &= A\frac{\partial \sigma(x)}{\partial x}dx \\ &= A\frac{\partial}{\partial x}\left(E\frac{\partial u(x,t)}{\partial x}\right)dx \\ &= AE\frac{\partial^2 u(x,t)}{\partial x^2}dx \end{aligned} \tag{7.3}$$

Putting this back into the original expression, we get:

$$\boxed{\frac{\partial^2 u(x,t)}{\partial t^2} = \frac{E}{\rho}\cdot\frac{\partial^2 u(x,t)}{\partial x^2}} \tag{7.4}$$

which is actually the famous wave-equation in classical wave mechanics! From this, we immediately see that the *wave velocity* (speed of wave propagation) is given by:

$$v_{wave} = c = \sqrt{\frac{E}{\rho}} \tag{7.5}$$

with general solutions to this equation taking the form $u(x,t) = f_1(x-ct) + f_2(x+ct)$.

§7.2 Explicit Solutions

To solve for explicit solutions to $u(x,t)$, we will consider the simple *separable* case. That is to say, we assert that the solution can be written as a product of purely single-variable functions as follows:

$$u(x,t) = A(x)\cdot B(t) \tag{7.6}$$

Plugging this into our wave equation, we get:

$$\begin{aligned} \frac{\partial^2}{\partial t^2}\left[A(x)\cdot B(t)\right] &= c^2\cdot\frac{\partial^2}{\partial x^2}\left[A(x)\cdot B(t)\right] \\ \Rightarrow\ \frac{1}{B(t)}\cdot\frac{\partial^2 B(t)}{\partial t^2} &= \frac{c^2}{A(x)}\cdot\frac{\partial^2 A(x)}{\partial x^2} \end{aligned} \tag{7.7}$$

which effectively separates out the variables into the left and right hand sides of the equation. They are thus related by a separation constant which we shall call $-c^2p^2$. With this, we now have the following 2 equations:

$$\frac{\partial^2 B(t)}{\partial t^2} = -c^2p^2B(t) \tag{7.8}$$

$$\frac{\partial^2 A(x)}{\partial x^2} = -p^2A(x) \tag{7.9}$$

These are in fact just the equations for a simple harmonic oscillator, which gives us purely sinusoidal solutions:

$$B(t) = b_p \sin(cpt) + \tilde{b}_p \cos(cpt) \tag{7.10}$$

$$A(x) = a_p \sin(px) + \tilde{a}_p \cos(px) \tag{7.11}$$

This assertion of separability may seem an unlikely scenario in real life, but what these ideal solutions allow us to do is construct arbitrary wave solutions by linear combination. This stems from the fact that the wave equation is a *linear equation*. As such, we have that general wave solutions can always take the form:

$$\boxed{u(x,t) = \sum_p \Big[a_p \sin(px) + \tilde{a}_p \cos(px)\Big] \cdot \Big[b_p \sin(cpt) + \tilde{b}_p \cos(cpt)\Big]} \tag{7.12}$$

where the $a_p, \tilde{a}_p, b_p\tilde{b}_p$ coefficients can all be found by looking at the *boundary conditions* and *initial conditions* of the system.

§7.3 Boundary Conditions

In this section, we will be exploring the behaviour of an elastic beam when it is subject to *boundary conditions*. Boundary conditions are constraints at the boundary of a system (material volume) that result in often unique and interesting system dynamics. The physical system subject to boundary conditions we will be studying is an elastic beam rigidly attached at one end.

§7.3.1 One-Sided Rigid Attachment

As mentioned, consider a beam rigidly attached to a surface at one end. Gravity is ignored for purposes of our analysis. For clarity, we will label the free end of the beam with letter A, and the rigidly attached end as B. This physical constraint imposes the following relation on our wave solution:

$$u(x = x_B, t) = 0 \tag{7.13}$$

Furthermore, since there are no forces acting on the end at A, we have that the strain at A should vanish:

$$\Rightarrow \quad A \cdot \frac{\partial}{\partial x} u(x = x_A, t) = 0 \tag{7.14}$$

Now consider if we perturb the beam such that we send a wave propagating through the material. We know all non-trivial solutions cannot have both the wave solution and its first spatial derivative vanish at the same point (by the theorem on *uniqueness*). Using this and the 2 boundary conditions asserted above, we see that propagating waves at both ends will be *reflected*, however there will only be a change of parity (flip) of the reflected wave at boundary B.

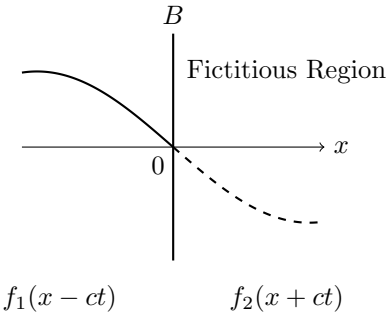

Figure 7.2: Solutions at Boundary B

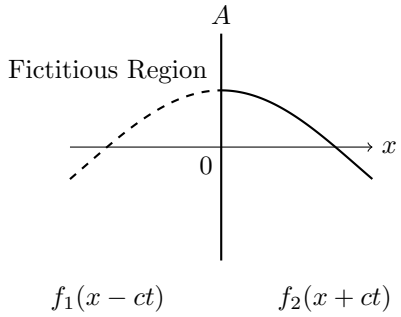

Figure 7.3: Solutions at Boundary A

The solutions near the boundaries are illustrated in figures 7.2 and 7.3 above, whereby the fictitious regions indicate regions past the boundary (outside the material volume). These are included for mathematical convenience. Note that the wave which propagate through the material are longitudinal (since $u(x,t)$ are x-directional displacements). Due to this parity flip upon reflection off boundary B, we see that the wave form experiences periodicity with a period of $T = \frac{4(x_B - x_A)}{c}$ (cycles over its configuration states).

From here, we now look to finding explicit forms for $u(x,t)$ when subject to these boundaries. Before we do this, we can simplify our calculations by defining $x_A = 0$ and $x_B = L$. Utilizing the separable solutions found earlier, we have:

$$\begin{aligned} & u(x=L,t) = 0 \\ \Rightarrow\ & a_p \sin(pL) + \tilde{a}_p \cos(pL) = 0 \end{aligned} \tag{7.15}$$

$$\begin{aligned} & \frac{\partial}{\partial x} u(x=0,t) = 0 \\ \Rightarrow\ & p \cdot a_p \cos(0) - p \cdot \tilde{a}_p \sin(0) = 0 \end{aligned} \tag{7.16}$$

$$\Rightarrow\ a_p = 0, \quad p = \frac{(2n+1)}{2L}\pi \tag{7.17}$$

where $n \in \mathbb{Z}$. With this result, we can write the general wave solution explicitly as follows:

$$\boxed{u(x,t) = \sum_p \left[\tilde{a}_p \cos\left(\frac{(2n+1)\pi}{2L}x\right)\right] \cdot \left[b_p \sin\left(\frac{(2n+1)\pi}{2L}ct\right) + \tilde{b}_p \cos\left(\frac{(2n+1)\pi}{2L}ct\right)\right]} \tag{7.18}$$

§7.4 Catenary Curvature

Let us now study effective 1D materials (strings) which are inelastic, but perfectly flexible (e.g. frictionless chains, ideal cables). The particular set-up we will be considering is such a string

attached at both ends, allowed to freely hang under the effect of gravity. The aim of this study is to determine the resulting shape of the string from **equilibrium** analysis of its differential segments. This physical system is known as a *catenary* and can be generalized to 2D membranes as well.

Consider first a differential segment of the string with length ds. We know intuitively that if there is some slack, the string will sag, causing its segments to tilt at some angle θ. This can be visualized as follows:

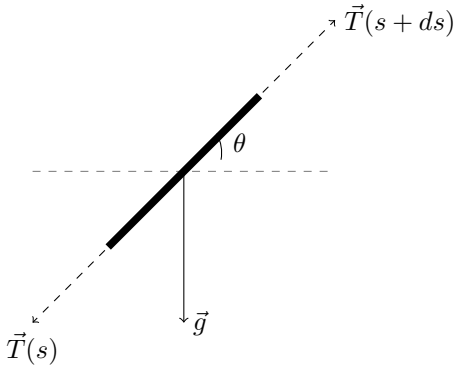

Figure 7.4: Differential Segment of String

Above, $\vec{T}(s)$ is the axial tensile force on the string a distance s along from the lowest point, and $\vec{g}$ is the force of gravity. Setting up a Cartesian coordinate system ($+x$ being horizontal and rightward, $+y$ being vertical and upward), we have that:

$$ds = \sqrt{1 + \left(\frac{dy}{dx}\right)^2}\, dx \tag{7.19}$$

At equilibrium, resolution of forces along the Cartesian axes gives us:

$$T\cos\theta = T_0, \quad T\sin\theta = \lambda g s \tag{7.20}$$

$$\Rightarrow \quad \tan\theta = \frac{dy}{dx} = \frac{\lambda g s}{T_0} \tag{7.21}$$

where T_0 is the tension at the lowest point C on the string and $\lambda g s$ constitutes the weight of string segment s away from C. Taking derivatives of these compoenents with respect to s, we get:

$$\frac{d}{ds}T\cos\theta = 0, \quad \frac{d}{ds}T\sin\theta = \lambda g \tag{7.22}$$

$$\Rightarrow \quad \frac{d}{ds}\tan\theta = \frac{\lambda g}{T_0} \tag{7.23}$$

$$\Rightarrow \quad \frac{1}{\sqrt{1 + \left(\frac{dy}{dx}\right)^2}} \frac{d}{dx}\left(\frac{dy}{dx}\right) = \frac{\lambda g}{T_0} \tag{7.24}$$

From this, we get a second order differential equation in x for the function $y(x)$:

$$\frac{d^2y}{dx^2} = \frac{\lambda g}{T_0}\sqrt{1+\left(\frac{dy}{dx}\right)^2} \tag{7.25}$$

By substituting $z = \frac{dy}{dx}$, we get that the solutions to this differential equation are hyperbolic cosine functions:

$$\boxed{y(x) = \frac{T_0}{\lambda g}\cosh\left(\frac{\lambda g}{T_0}x + \alpha\right) + \beta} \tag{7.26}$$

where α and β are geometry dependent constants determined by the relative locations of the pins from the defined Cartesian coordinates. In this analysis, we arrived at cosh functions since they are symmetric about the origin, for which we defined the lowest position of the string to have coordinate $x = 0$. It is well known that our system is mirror symmetric about this points about the y-axis.

Chapter 8

3D Rigid Body Dynamics

Earlier on in the course, we have looked at rigid body systems whereby the mass distribution, or moment of inertia of these bodies remains time-invariant. In this chapter, we drop this assumption, which extends our analysis of such dynamical systems. We will also be looking at an interesting new way of approaching dynamical analysis of a 3D rigid body system through the use of Euler angles.

§8.1 Rigid-Body Coordinates

Recall that the angular momentum of a rigid body is defined as $\vec{H} = I\vec{\omega}$, where I is the inertia tensor. If we now take the time derivative of this, we get:

$$\dot{\vec{H}} = \frac{d}{dt}(I)\omega + I\frac{d}{dt}(\omega) \tag{8.1}$$

The difficulty of approaching such a calculation lies in the time derivative of the inertia tensor. To simplify our analysis, we can instead consider some coordinate system that we fix to the rigid body, which will effectively allow us to utilize principal axes properties. First we note that we can write the momenta of forces as:

$$\begin{aligned} \dot{\vec{H}} &= \dot{\vec{H}}_r + \vec{\omega} \times \vec{H} \\ &= I\dot{\vec{\omega}} + \vec{\omega} \times (I\vec{\omega}) \end{aligned} \tag{8.2}$$

The goal now is to be able to perform a coordinate transformation such that we move into a reference frame with our inertia tensor being purely diagonal (principal axis). Explicitly, we want some map that takes:

$$M_i = \dot{H}_i - \varepsilon_{ijk}(H_j\omega_k - H_k\omega_j) \quad \rightarrow \quad M_i = I_i\dot{\omega}_i - \varepsilon_{ijk}(I_j - I_k)\omega_j\omega_k$$

The principal axis reference frame is ideal as we will have that:

$$M_i = I_i\dot{\omega}_i - \varepsilon_{ijk}(I_j - I_k)\omega_j\omega_k = 0 \tag{8.3}$$

Before we continue, it would first be useful to introduce the concept of *Euler angles* and Euler's rotation theorem.

Theorem 8.1.1. *Any arbitrary rotation in 3-dimensional Euclidean space can be parameterized by just 3 real valued parameters.*

From the theorem above, we can actually represent any rotation via the product of 3 consecutive matrices, each defined by a single parameter (making up the 3 rotation parameters). Since we are dealing with rigid rotations, these parameters will naturally be angular coordinates, which we denote as ϕ, θ and ψ (Euler angles). To define these angles, consider an unprimed ($\{x, y, z\}$) and a primed ($\{x', y', z'\}$) coordinate system. The unprimed coordinates denote the initial state, whereas the primed the rotated state.

Note: Since we are only dealing only with rotations, the origin of the primed and unprimed coordinate systems remain coincident.

The Euler angles are thus defined (with their associated rotation matrices) as follows:

1. ϕ: *The first rotation about the z-axis.*

$$R_1(\phi) = \begin{bmatrix} \cos\phi & \sin\phi & 0 \\ -\sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{8.4}$$

2. θ: *The second rotation about the current x'-axis ($\theta \in [0, \pi]$).*

$$R_2(\theta) = \begin{bmatrix} \cos\theta & 0 & \sin\theta \\ 0 & 1 & 0 \\ -\sin\theta & 0 & \cos\theta \end{bmatrix} \tag{8.5}$$

3. ψ: *The third rotation about the current z'-axis.*

$$R_3(\psi) = \begin{bmatrix} \cos\psi & \sin\psi & 0 \\ -\sin\psi & \cos\psi & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{8.6}$$

In order to accurately produce the given rotation via this 3 parameter decomposition, the matrices must be applied to the initial coordinates in the correct sequence (i.e. left multiply R_1 then R_2 then R_3). Explicitly, we write this as:

$$\begin{aligned} \begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} &= R_3(\psi)R_2(\theta)R_1(\phi) \begin{bmatrix} x \\ y \\ z \end{bmatrix} \\ &= \begin{bmatrix} \cos\psi & \sin\psi & 0 \\ -\sin\psi & \cos\psi & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \cos\theta & 0 & \sin\theta \\ 0 & 1 & 0 \\ -\sin\theta & 0 & \cos\theta \end{bmatrix} \begin{bmatrix} \cos\phi & \sin\phi & 0 \\ -\sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} \end{aligned} \tag{8.7}$$

The progressive rotations can be visualized as in figure 8.1 below.

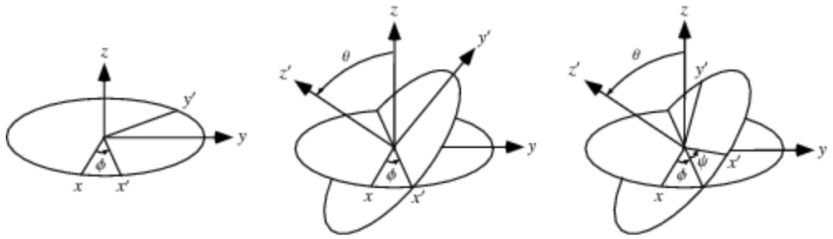

Figure 8.1: Progressive Rotations by Euler Angles

Due to the way in which we have defined the rotations, the appropriate unit vectors to use would be the set of $\{\hat{k}, \hat{i}', \hat{k}'\}$, where k is the direction of the z-axis, i' the x'-axis and k' the z'-axis. Using this, we know that the angular velocity of our rigid body in terms of time-derivatives of the Euler angles is:

$$\vec{\omega} = \dot{\phi}\hat{k} + \dot{\theta}\hat{i}' + \dot{\psi}\hat{k}' \tag{8.8}$$

At the same time, we also know that via the rotation matrices applied to the original $\{\hat{i}, \hat{j}, \hat{k}\}$ unit vectors, we have:

$$\begin{aligned}
\vec{\omega} &= \dot{\phi}\left(\sin\psi\cos\theta\hat{i} + \cos\psi\sin\theta\hat{j} + \cos\theta\hat{k}\right) + \dot{\theta}\left(\cos\psi\hat{i} - \sin\psi\hat{j}\right) + \dot{\psi}\hat{k} \\
&= \left(\dot{\phi}\sin\psi\sin\theta + \dot{\theta}\cos\psi\right)\hat{i} + \left(\dot{\phi}\cos\psi\sin\theta - \dot{\theta}\sin\psi\right)\hat{j} + \left(\dot{\phi}\cos\theta + \dot{\psi}\right)\hat{k} \\
&= \begin{bmatrix} \dot{\phi}\sin\psi\sin\theta + \dot{\theta}\cos\psi \\ \dot{\phi}\cos\psi\sin\theta - \dot{\theta}\sin\psi \\ \dot{\phi}\cos\theta + \dot{\psi} \end{bmatrix}
\end{aligned} \tag{8.9}$$

From here, what we can do is project the angular momentum onto its Cartesian coordinate components (in angular variables), which will allow us to construct a set of coupled differential equations for the Euler angles. This is done as follows:

$$\begin{aligned}
\vec{H} &= H\left(\sin\theta\sin\psi\hat{i} + \sin\theta\cos\psi\hat{j} + \cos\theta\hat{k}\right) \\
&= \left(\dot{\phi}\sin\psi\sin\theta + \dot{\theta}\cos\psi\right)\hat{i} + \left(\dot{\phi}\cos\psi\sin\theta - \dot{\theta}\sin\psi\right)\hat{j} + \left(\dot{\phi}\cos\theta + \dot{\psi}\right)\hat{k}
\end{aligned} \tag{8.10}$$

$$\begin{aligned}
\Rightarrow\ & H\sin\theta\sin\psi = I_x\left(\dot{\phi}\sin\psi\sin\theta + \dot{\theta}\cos\psi\right) \\
& H\sin\theta\cos\psi = I_y\left(\dot{\phi}\cos\psi\sin\theta - \dot{\theta}\sin\psi\right) \\
& H\cos\theta = I_z\left(\dot{\phi}\cos\theta + \dot{\psi}\right)
\end{aligned} \tag{8.11}$$

Chapter 9

Hamiltonian Mechanics

The final chapter will introduce yet another formalism of classical mechanics. Here, we utilize the ***Hamiltonian*** *to analyze the mechanics of a dynamical system (instead of the Lagrangian). The Hamiltonian formalism is extremely useful in the study of quantum mechanics, and so gives a natural segue into further studies of physics. We begin by building an intuition of Hamiltonians from Lagrangian mechanics which we have now familiarized ourselves with. This will be done via the* ***Legendre transformations****.*

§9.1 Legendre Transformations

Derived by Adrien-Marie Legendre, the Legendre transform is an *involutive* transformation on the real-valued convex functions of one real variable. Simply put, it converts one set of conjugate variables to another.

Definition 9.1.1. Convex Function: *Given a function $f(x)$ over a domain X, we say that $f(x)$ is convex if $\forall x_1, x_2 \in X$ and $\forall t \in [0,1]$, we have that*

$$f(t \cdot x_1 + (1-t) \cdot x_2) \leq t \cdot f(x_1) + (1-t) \cdot f(x_2) \tag{9.1}$$

$f(x)$ is ***strictly convex*** *if the equality is dropped in (9.1) .*

In classical mechanics, we use this transformation to move between the Lagrangian and Hamiltonian and can be thought of as analogous to moving between the time and frequency domains with Fourier transforms.

Definition 9.1.2. Legendre Transform: *Given a strictly convex and differentiable function $f : \mathbb{R} \to \mathbb{R}$, the Legendre transform $G(p) = \mathbb{L}\{f(x)\}(p)$ is defined as*

$$\mathbb{L}\{f(x)\}(p) = \max_x \{xp - f(x)\} \tag{9.2}$$

$$\text{or} \quad G(p) = x_p p - f(x_p) \tag{9.3}$$

Where in (9.3), x_p solves the relation $p = f'(x_p)$. Alternative, we can rewrite (9.3) to reverse the independent and dependent variables, which gives:

$$f\big(x_p(p)\big) + G(p) = f'\big(x_p(s)\big) \cdot x_p(p) \tag{9.4}$$

This way of writing the definition allows use to use a geometric approach to understand what the Legendre transform is. The Legendre transform is the difference between the height of the tangent and the function value which shows a geometric duality transformation. It is useful now to present some properties and proofs of the Legendre transform. **Properties:**

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

   Proof. Let $f(x)$ be a strictly convex and differentiable function, then its Legendre transform is given as
   $$G(p) = x(p)p - f\big(x(p)\big) \tag{9.5}$$
   Then considering the derivative of $G(p)$ with respect to p,
   $$\begin{aligned} \frac{dG(p)}{dp} &= \frac{d}{dp}\Big(x(p)p\Big) - \frac{d}{dp} f\big(x(p)\big) && (9.6)\\ &= x + x'(p)p - f'\big(x(p)\big)x'(p) && (9.7)\\ &= x + x'(p)p - x'(p)p && (9.8)\\ &= x && (9.9) \end{aligned}$$
   Now considering the second derivative,
   $$\begin{aligned} \frac{d^2G(p)}{dp^2} &= x'(p) = \frac{1}{p'(x)} && (9.10)\\ &= \frac{1}{f''(x)} > 0 && (9.11) \end{aligned}$$
   since $f(x)$ is already convex, which thus shows that G(p) also fulfills the second derivative test and is indeed strictly convex. □

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

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

   Proof. Let $f(x)$ be a convex differentiable function, then by definition,
   $$\begin{aligned} & G(p) = \mathbb{L}\{f(x)\}(p) = x(p)f'\big(x(p)\big) - f\big(x(p)\big) \\ \Rightarrow\ & G\big(G(p)\big) = \mathbb{L}\{\mathbb{L}\{f(x)\}\} = pG'(p) - G(p) \end{aligned}$$

Then since

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

Showing that the Legendre transform is indeed an involution. □

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

$$\begin{aligned} & \mathcal{L}(q, \dot{q}) = \frac{1}{2} m \dot{q}^2 - v(q) \\ \Rightarrow \; & G(p) = \mathbb{L}\{\mathcal{L}\big(q, \dot{q}(p)\big)\}(p) = p\dot{q}(p) - \mathcal{L}\big(q, \dot{q}(p)\big) \end{aligned} \tag{9.12}$$

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

$$\begin{aligned} & \begin{cases} x \to \dot{q} \\ p \to p \end{cases} \\ \Rightarrow \; & p = \frac{\partial \mathcal{L}}{\partial \dot{q}} = m\dot{q} \\ \Rightarrow \; & \dot{q} = \frac{p}{m} \quad \Rightarrow \quad \mathcal{L}\big(q, \dot{q}(p)\big) = \frac{p^2}{2m} - V(q) \\ \Rightarrow \; & G(p) = \frac{p^2}{m} - \frac{p^2}{2m} + V(q) = \frac{p^2}{2m} + V(q) \end{aligned} \tag{9.13}$$

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

$$\boxed{\mathcal{H} = \frac{p^2}{2m} + V(q)} \tag{9.14}$$

which is exactly $T + V$, the total energy of the system! However, the Hamiltonian is **not** always the energy of the system, like for instance when $T = T(q, \dot{q})$. With this transformation, we have gone from a set of $\{q_j, \dot{q}_j\}$ variables to a set of $\{q_j, p_j\}$ variables. Extending this to a general N degree of freedom system, we can write the Hamiltonian as:

$$\boxed{H(\vec{p}, \vec{q}) = \sum_j p_j \dot{q}_j(\vec{p}) - \mathcal{L}\left(\vec{q}, \dot{\vec{q}}(\vec{p})\right)} \tag{9.15}$$

§9.2 Hamilton's Equations

The next question is how do we retrieve the dynamics of a system from the Hamiltonian? The answer lies in *Hamilton's equations* which are the Hamiltonian equivalent of the Euler-Lagrange

equations. How we arrive at these are as follows. Consider the differential of a Lagrangian:

$$\begin{aligned} d\mathcal{L} &= \sum_j \frac{\partial \mathcal{L}}{\partial q_j} dq_j + \sum_j \frac{\partial \mathcal{L}}{\partial \dot{q}_j} d\dot{q}_j \\ &= \sum_j \dot{p}_j dq_j + \sum_j p_j d\dot{q}_j \end{aligned} \tag{9.16}$$

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

$$\sum_j p_j d\dot{q}_j = d\left(\sum_j p_j \dot{q}_j \right) - \sum_j \dot{q}_j dp_j \tag{9.17}$$

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

$$\Rightarrow \; \boxed{\dot{q}_j = \frac{\partial \mathcal{H}}{\partial p_j}, \;\; \dot{p}_j = -\frac{\partial \mathcal{H}}{\partial q_j}} \tag{9.18}$$

Where the equations in (9.18) are the Hamilton's equations (a.k.a. *canonical equations*). These can be used to derive the equations of motion of a system just as the Euler-Lagrange equations would. Notice that these are first order differential equations and are usually much easier to solve analytically, but the complexity of Hamiltonian mechanics arises in the formulation of the Hamiltonian itself. Hamiltonians are usually difficult to just write down and are most easily found from the Lagrangian.

What is useful about using the Hamiltonian rather than Lagrangian formalism is that if we have a Hamiltonian that is not an explicit function of time, it is **always** conserved! This holds **even** for system where the energy is not.

Theorem 9.2.1. *Given a system such that its Hamiltonian is not an explicit function of time, then it follows that the Hamiltonian is conserved.*

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

$$\frac{d\mathcal{H}}{dt} = \frac{\partial \mathcal{H}}{\partial t} + \sum_j \dot{p}_j \frac{\partial \mathcal{H}}{\partial p_j} + \sum_j \dot{q}_j \frac{\partial \mathcal{H}}{\partial q_j} \tag{9.19}$$

$$= \frac{\partial \mathcal{H}}{\partial t} + \sum_j p_j \dot{q}_j - \sum_j \dot{q}_j p_j = \frac{\partial \mathcal{H}}{\partial t} \tag{9.20}$$

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

§9.3 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:

$$\begin{aligned} \frac{d}{dt}f(p,q,t) &= \frac{\partial f}{\partial t} + \dot{q}\frac{\partial f}{\partial q} + \dot{p}\frac{\partial f}{\partial p} \\ &= \frac{\partial f}{\partial t} + \frac{\partial \mathcal{H}}{\partial p}\frac{\partial f}{\partial q} - \frac{\partial \mathcal{H}}{\partial q}\frac{\partial f}{\partial p} \\ \Rightarrow \quad \frac{d}{dt}f - \frac{\partial}{\partial t}f &= \frac{\partial \mathcal{H}}{\partial p}\frac{\partial f}{\partial q} - \frac{\partial \mathcal{H}}{\partial q}\frac{\partial f}{\partial p} \equiv [\mathcal{H}, f] \end{aligned} \tag{9.21}$$

where $[\mathcal{H}, f]$ is known as the Poisson bracket of the Hamiltonian $\mathcal{H}$ with f. From this, we see that if the function f we are working with is **not** an explicit function of time, (9.21) reduces to

$$\boxed{\frac{d}{dt}f(p,q) = [\mathcal{H}, f]} \tag{9.22}$$

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 $f = f(t)$.

$$\frac{d}{dt}f(t) = \frac{\partial}{\partial t}f(t) \tag{9.23}$$

$$\Rightarrow \quad [\mathcal{H}, f] = \frac{d}{dt}f(t) - \frac{\partial}{\partial t}f(t) = 0 \tag{9.24}$$

But in this scenario, this does **not** imply that f is conserved. Equation (9.21) actually gave a specific example of a Poisson bracket. We can generalize the definition as follows.

Definition 9.3.1. Poisson Bracket: *Given 2 functions of the conjugate variables $f(\vec{p}, \vec{q}, t)$ and $g(\vec{p}, \vec{q}, t)$, their Poisson bracket is defined as*

$$[f, g]_{(p,q)} \equiv \sum_j \left(\frac{\partial f}{\partial p_j}\frac{\partial g}{\partial q_j} - \frac{\partial f}{\partial q_j}\frac{\partial g}{\partial p_j} \right) \tag{9.25}$$

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

Properties:

1. *Anti-symmetric*: $[f, g] = -[g, f]$
2. *Distributive*: $[f + g, h] = [f, h] + [g, h]$

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

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

Theorem 9.3.1. *Given 2 integrals of motion f and g that are not explicit functions of time i.e.*

$$\frac{df}{dt} = \frac{\partial f}{\partial t} = 0, \quad \frac{dg}{dt} = \frac{\partial g}{\partial t} = 0 \tag{9.26}$$

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

$$\Rightarrow \frac{d}{dt}[f, g] = 0 \tag{9.27}$$

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

$$[\mathcal{H}, f] = [\mathcal{H}, g] = 0 \tag{9.28}$$

Then utilizing the Jacobi identity, we get

$$[\mathcal{H}, [f, g]] + [f, [g, \mathcal{H}]] + [g, [\mathcal{H}, f]] = 0 \tag{9.29}$$

$$\Rightarrow \; [\mathcal{H}, [f, g]] = 0 \tag{9.30}$$

$$\Rightarrow \; \frac{d}{dt}[f, g] = 0 \tag{9.31}$$

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

It is now good to establish, just as we have with Lagrangian mechanics, a systematic work-flow for Hamiltonian mechanics problems. The sequence of steps are as follows.

Hamiltonian Work-Flow:

1. Choose a set of generalized coordinates $\{q_j, \dot{q}_j\}$ best suited for the problem.
2. Write the kinetic energy, potential energy and thus Lagrangian ($T(\vec{q}, \dot{\vec{q}})$, $V(\vec{q})$ and $\mathcal{L}(\vec{q}, \dot{\vec{q}})$) expressed in terms of the chosen generalized coordinates.
3. Compute the conjugate momenta $p_j = \frac{\partial \mathcal{L}}{\partial \dot{q}_j}$ of the system.

4. Formulate the Hamiltonian $H(\vec{p},\vec{q}) = \sum_j p_j \dot{q}_j(\vec{p}) - \mathcal{L}\left(\vec{q},\dot{\vec{q}}(\vec{p})\right)$.
5. Utilize the Hamilton's equations of motion (9.18) to construct the equations of motion of the system.

Looking through the steps above, it seems rather redundant to utilize the Hamiltonian approach if we first have to formulate the Lagrangian anyway. Is there then a means to construct the Hamiltonian without the Lagrangian? Turns out the answer is yes! These are primarily done in **2 methods**.

1. "*Guess and Check*"

   If the mechanical system we are dealing with is simple enough, it is often possible to determine its generalized coordinates and momenta from simply eyeballing it. Of course we cannot just stop there and blindly work with these guesses, we need a way to verify them.

   To do so, we utilize the **Poisson bracket** properties. That is, any canonical set of variables chosen for a problem **must** satisfy the *canonical relations*. Presenting this explicitly, consider first having a set of canonical variables $\{q_j, p_j\}$ that we want to transform to another set of canonical variables $\{Q_j, P_j\}$. We assume that the original set $\{q_j, p_j\}$ satisfies the canonical relations. We then also require our new set $\{Q_j, P_j\}$ to do the same which means,

   $$[Q_j, Q_k]_{p,q} = 0, \quad [P_j, P_k]_{p,q} = 0, \quad [P_j, Q_k]_{p,q} = \delta_{j,k} \tag{9.32}$$

   where we have performed these Poisson brackets with respect to the old set of canonical variables $\{q_j, p_j\}$.

2. *Generator Functions*

   This next approach is far more mathematically rigorous and consistent than the previous one, making it the more widely adopted method of Hamiltonian formulation (refer to the following section for more details).

§9.4 Generator Functions

The motivation behind this goes back to a property of the Lagrangian, whereby adding a total time derivative term to a Lagrangian does **not** change the physics (resultant equations of motions). To exploit this, we have to use the *principle of least action*.

Principal of Least Action

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

$$S = \int_{t_1}^{t_2} \mathcal{L}(q, \dot{q}, t) dt \tag{9.33}$$

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

From this principle, we see that the variation of the action must vanish on the trajectory that minimizes the action:

$$\Rightarrow \delta \int_{t_1}^{t_2} \mathcal{L} dt = \delta \int_{t_1}^{t_2} \Big(\sum_j p_j \dot{q}_j - \mathcal{H}(\vec{p}, \vec{q}, t) \Big) dt = 0 \tag{9.34}$$

Also say that we want to go from one set of canonical variables to another, $(\vec{p}, \vec{q}) \to (\vec{P}, \vec{Q})$, which would cause our Hamiltonian to transform as $\mathcal{H}(\vec{p}, \vec{q}, t) \to \mathcal{K}(\vec{P}, \vec{Q}, t)$. As such, our variational principle can be rewritten as

$$\delta \int_{t_1}^{t_2} \Big(\sum_j P_j \dot{Q}_j - \mathcal{K}(\vec{P}, \vec{Q}, t) + \frac{d}{dt} F(\vec{q}, \vec{p}, \vec{Q}, \vec{P}, t) \Big) dt = 0 \tag{9.35}$$

where we have exploited the Lagrangian property mentioned above for some arbitrary function $F(\vec{q}, \vec{p}, \vec{Q}, \vec{P}, t)$, which is known as a *generator function*. Let us first consider a particular case where $F = F(\vec{q}, \vec{Q}, t)$.

$$dF(\vec{q}, \vec{Q}, t) = \frac{\partial F}{\partial t} + \sum_j \frac{\partial F}{\partial q_j} dq_j + \sum_j \frac{\partial F}{\partial Q_j} dQ_j \tag{9.36}$$

$$\Rightarrow \frac{d}{dt} F(\vec{q}, \vec{Q}, t) = \frac{\partial F}{\partial t} + \sum_j \dot{q}_j \frac{\partial F}{\partial q_j} + \sum_j \dot{Q}_j \frac{\partial F}{\partial Q_j} \tag{9.37}$$

$$\begin{aligned} \Rightarrow \sum_j p_j \dot{q}_j - \mathcal{H} &= \sum_j P_j \dot{Q}_j - \mathcal{K} + \left(\frac{\partial F}{\partial t} + \sum_j \dot{q}_j \frac{\partial F}{\partial q_j} + \sum_j \dot{Q}_j \frac{\partial F}{\partial Q_j} \right) \\ &= \Big(\sum_j P_j - \frac{\partial F}{\partial Q_j} \Big) \dot{Q}_j + \sum_j \frac{\partial F}{\partial q_j} \dot{q}_j - \Big(\mathcal{K} - \frac{\partial F}{\partial t} \Big) \end{aligned} \tag{9.38}$$

If we compare the right and left-hand sides of equation (9.38), we see that we get definitive relations between the canonical coordinates. These relations are

$$\sum_j \Big(P_j - \frac{\partial F}{\partial Q_j} \Big) = 0, \quad p_j = \frac{\partial F}{\partial q_j}, \quad \mathcal{H} = \mathcal{K} - \frac{\partial F}{\partial t} \tag{9.39}$$

This shows that choosing some function of the form $F(\vec{q}, \vec{Q}, t)$ allows use to 'generate' relations between the canonical variables, which would then allow us to construct our Hamiltonian! These coordinate transformations are known as *point transformations*. In fact, there are 4 different forms of generator functions with different associated point transformations. The one we have just done is known as a *generator function of the 1^{st} kind*. We will list all of them and their point transformations below.

1. *Generator Functions of the 1^{st} Kind*:

$$F_1 = F(\vec{q}, \vec{Q}, t) \tag{9.40}$$

$$\Rightarrow P_j = \frac{\partial F_1}{\partial Q_j}, \quad p_j = \frac{\partial F_1}{\partial q_j}, \quad \mathcal{K} = \mathcal{H} + \frac{\partial F_1}{\partial t} \tag{9.41}$$

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

$$F_2 = F(\vec{q}, \vec{P}, t) - \vec{Q} \cdot \vec{P} \tag{9.42}$$

$$\Rightarrow \; Q_j = \frac{\partial F_2}{\partial P_j}, \;\; p_j = \frac{\partial F_2}{\partial q_j}, \;\; \mathcal{H} = \mathcal{K} + \frac{\partial F_2}{\partial t} \tag{9.43}$$

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

$$F_3 = F(\vec{P}, \vec{Q}, t) + \vec{q} \cdot \vec{p} \tag{9.44}$$

$$\Rightarrow \; q_j = -\frac{\partial F_3}{\partial P_j}, \;\; P_j = -\frac{\partial F_3}{\partial Q_j}, \;\; \mathcal{H} = \mathcal{K} + \frac{\partial F_3}{\partial t} \tag{9.45}$$

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

$$F_4 = F(\vec{p}, \vec{P}, t) - \vec{Q} \cdot \vec{P} + \vec{q} \cdot \vec{p} \tag{9.46}$$

$$\Rightarrow \; q_j = -\frac{\partial F_4}{\partial p_j}, \;\; Q_j = \frac{\partial F_4}{\partial P_j}, \;\; \mathcal{H} = \mathcal{K} + \frac{\partial F_4}{\partial t} \tag{9.47}$$

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.

§9.5 Hamilton-Jacobi Equation

Despite appearances, the very fact that we have the freedom to choose your generator functions actually make them a powerful tool! To see this, consider a 1 degree of freedom system, having the canonical variables $\{p, q\}$. Then let's say we want to perform some point transformation on these variables to another set of canonical variables such that

$$\{p, q\} \to \{P, Q\} \tag{9.48}$$

$$\mathcal{H}(p, q, t) \to \mathcal{K}(P, Q, t) \tag{9.49}$$

Then using the generator function relation of the Legendre transforms

$$p\dot{q} - \mathcal{H}(p, q, t) = P\dot{Q} - \mathcal{K}(P, Q, t) + \frac{d}{dt}F(p, q, P, Q, t) \tag{9.50}$$

If we pick specifically a generator function of the 2^{nd} kind, this would give the point transforms,

$$F_2 = F(q, P, t) - QP \tag{9.51}$$

$$\Rightarrow \; Q = \frac{\partial F_2}{\partial P}, \;\; p = \frac{\partial F_2}{\partial q}, \;\; \mathcal{K} = \mathcal{H} + \frac{\partial}{\partial t}F \tag{9.52}$$

From here, due to our liberty to choose any generator function of our choice that includes a subset of all the canonical variables, let us choose F such that

$$\mathcal{K}(P, Q, t) = \mathcal{H}(p, q, t) + \frac{\partial}{\partial t}F(p, q, P, Q, t) = 0 \tag{9.53}$$

An immediate result of this choice guarantees that our new canonical variables $\{P, Q\}$ are **conserved quantities**, since

$$\dot{P} = -\frac{\partial \mathcal{K}}{\partial Q} = 0, \quad \dot{Q} = \frac{\partial \mathcal{K}}{\partial P} = 0 \tag{9.54}$$

keeping in mind that P and Q are constants, then we can write the generator as

$$F(q, P, t) = S(q, t) + A \tag{9.55}$$

where A is some constant term. Finally, since

$$p = \frac{\partial F_2}{\partial q} = \frac{\partial S}{\partial q}, \quad Q = \frac{\partial S}{\partial P} \tag{9.56}$$

$$\Rightarrow \quad \boxed{\mathcal{H}(q, \frac{\partial s}{\partial q}, t) + \frac{\partial S}{\partial t} = 0} \tag{9.57}$$

where (9.57) is known as the *Hamilton-Jacobi equation*. As mentioned earlier, what is good about picking a generator that satisfies this form is that it allows us to **always** retrieve canonical variables which are conserved quantites.

It was not by random selection that we have chosen S as the name of our generator function above. To see why we have done so, consider the following.

$$\begin{aligned} & S = S(q, t) \\ \Rightarrow \quad & \frac{dS}{dt} = \dot{q}\frac{\partial S}{\partial q} + \frac{\partial S}{\partial t} = \dot{q}p - \mathcal{H}(p, q, t) \end{aligned}$$

where we used (9.56) and (9.57). Then by the definition of a Legendre transformation,

$$\begin{aligned} & \frac{dS}{dt} = \mathcal{L}(q, \dot{q}, t) \\ \Rightarrow \quad & S(q, t) = \int_{t_1}^{t_2} \mathcal{L}(q, \dot{q}, t) dt \end{aligned} \tag{9.58}$$

which is exactly the definition of the Action (9.33)!

§9.6 Infinitesimal Time Translations

In this section, we will look at how the Hamiltonian is itself a special generator function which generates small *time translations*. The idea behind this is that if we were to use the Hamiltonian as the generator of point transformations, the new canonical variables would simply be the old canonical variables shifted by a small amount in time! To show this, we first look the generator functions of the 2^{nd} kind. We can actually write them in a different way from how it was previously presented, and we will denote this alternate form with a prime.

$$F_2' = \vec{q} \cdot \vec{P} + \epsilon G(\vec{q}, \vec{P}) \tag{9.59}$$

Its associated point transformations are then

$$p_j = \frac{\partial F_2'}{\partial q_j} = P_j + \epsilon \frac{\partial G}{\partial q} \tag{9.60}$$

$$Q_j = \frac{\partial F_2'}{\partial P_j} = q_j + \epsilon \frac{\partial G}{\partial P_j} \tag{9.61}$$

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

$$\lim_{\epsilon \to 0} \frac{\partial G(\vec{q}, \vec{P})}{\partial P_j} = \frac{\partial G(\vec{q}, \vec{p})}{\partial p_j} \tag{9.62}$$

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

$$\boxed{P_j = p_j - dt \frac{\partial \mathcal{H}}{\partial q_j}, \quad Q_j = q_j + dt \frac{\partial \mathcal{H}}{\partial P_j}} \tag{9.63}$$

$$\Rightarrow \quad P_j = p_j + \dot{p}_j dt, \quad Q_j = q_j + \dot{q}_j dt$$

$$\Rightarrow \quad P_j(t) = p_j(t + dt), \quad Q_j(t) = q_j(t + dt) \tag{9.64}$$

Hence, we see from (9.63) and (9.64) that the Hamiltonian is indeed the *generator of time translation.*

§9.7 Phase Space

In the analysis of a dynamical systems, we now know how to assess the degrees of freedom of a system and to represent them as generalized coordinates. We also know how to perform point transformations on these coordinates with the aid of generator functions into new sets of canonical variables. It is only natural to then ask, is there a plot we could construct with these abstract coordinates to give us visual insights into our system? The answer to this lies in what is known as *phase space diagrams.*

Definition 9.7.1. Phase Space: *In a mechanical system, phase space is a multidimensional space in which every state of the system and their trajectories over time are represented. Every canonical momentum and position variable is an axis in the phase diagram.*

Phase diagrams are most easily read if the system can be represented by one canonical position and momentum pair. This is known as a *phase plane.* An example of this would be a simple pendulum system undergoing small oscillations.

A useful property of phase space is that all phase space trajectories do **not** intersect. This is due to the fact that every conjugate variable pair $\{q_j, p_j\}$ has a **unique** corresponding $\{\dot{q}_j, \dot{p}_j\}$ pair.

§9.7.1 Liouville's Theorem

An important aspect of looking at the phase space of a system is that it adheres to Liouville's theorem. To appreciate this theorem, we first need to know the definition of a the *phase space distribution function*.

Definition 9.7.2. *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 9.7.1. *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 9.7.2. *Given some vector field $\vec{F}$ and some region of space with volume V and boundary ∂V, the volume integral of the divergence of $\vec{F}$ and the surface integral of the field are related as follows:*

$$\int_V \boldsymbol{\nabla} \cdot \vec{F} dV = \int_{\partial V} \vec{F} \cdot d\vec{A} \tag{9.65}$$

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:

$$\begin{bmatrix} dq \\ dp \end{bmatrix} = \begin{bmatrix} \dot{q} \\ \dot{p} \end{bmatrix} dt = \vec{v} dt \tag{9.66}$$

where $\vec{v}$ 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:

$$dA = \int_C dl(\vec{n} \cdot \vec{v}) dt \tag{9.67}$$

where $\vec{n}$ 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:

$$\begin{aligned}\frac{dA}{dt} &= \int_C dl(\vec{n}\cdot\vec{v}) \\ &= \int_A \boldsymbol{\nabla}\cdot\vec{v}dA \\ &= \int_A dA\left(\frac{\partial}{\partial q}\dot{q}+\frac{\partial}{\partial p}\dot{p}\right) \\ &= \int_A dA\left(\frac{\partial}{\partial q}\frac{\partial\mathcal{H}}{\partial p}-\frac{\partial}{\partial p}\frac{\partial\mathcal{H}}{\partial q}\right)\end{aligned} \tag{9.68}$$

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:

$$\boxed{\frac{d}{dt}\rho(q_j,p_j)=\frac{\partial\rho}{\partial t}+\sum_j\left(\frac{\partial\rho}{\partial q_j}\dot{q}_j+\frac{\partial\rho}{\partial p_j}\dot{p}_j\right)=0} \tag{9.69}$$

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.