CLASSICAL MECHANICS

Prabuddha Chakraborty, notes by Ramdas Singh

Third Semester

List of Symbols

Placeholder

Contents

Chapter 1

THE LAWS OF MOTION AND CONSERVATION

1.1 An Introduction

1.1.1 Standard Conventions

July 21st.

We will be using the convention of $\hat{e}_x \times \hat{e}_y = \hat{e}_z$, where $\times$ denotes the *cross product*, and $\hat{e}_w$ is the unit vector along the w axis. For two general vectors $\vec{A}$ and $\vec{B}$, the i^{th} component of the cross product $\vec{C} = \vec{A} \times \vec{B}$ is given as

$$[\vec{C}]_i = \sum_{j,k} \epsilon_{ijk} [\vec{A}]_j [\vec{B}]_k, \tag{1.1}$$

where ϵ_{ijk} is the *Levi-Civita symbol*, which is simply the sign of the permutation of the indices ijk. A *dot product* between the two vectors is defined as

$$\vec{A} \cdot \vec{B} = \sum_i A_i B_i. \tag{1.2}$$

Along with the Levi-Civita symbol, we will also be using the *Kronecker delta* δ_{ij}, which is defined as $\delta_{ij} = 1$ if $i = j$ and 0 otherwise. It can be shown that

$$\epsilon_{ijk}\epsilon_{ilm} = \delta_{jl}\delta_{km} - \delta_{jm}\delta_{kl}. \tag{1.3}$$

Example 1.1. The following are some examples of the above notation:

- *Angular momentum*: $\vec{L} = \vec{r} \times \vec{p} \implies L_i = \epsilon_{ijk} r_j p_k$ where $\vec{p}$ is the momentum vector and $\vec{r}$ is the position vector.
- *Kinetic energy*: $T = \frac{1}{2} m \vec{v} \cdot \vec{v} = \sum_i \frac{1}{2} m v_i v_i = \sum_i \frac{p_i p_i}{2m}$, where $\vec{v}$ is the velocity vector.
- *Torque*: $\vec{\tau} = \vec{r} \times \vec{F} \implies \tau_i = \epsilon_{ijk} r_j F_k$, where $\vec{F}$ is the force vector.

Gradient, Divergence, and Curl

Definition 1.2. The *gradient* of a scalar field f is defined as

$$\vec{\nabla} f := \frac{\partial f}{\partial x} \hat{e}_x + \frac{\partial f}{\partial y} \hat{e}_y + \frac{\partial f}{\partial z} \hat{e}_z. \tag{1.4}$$

The gradient points in the direction of the steepest ascent of the function f. The gradient operator $\vec{\nabla}$

is also defined as

$$\vec{\nabla} := \frac{\partial}{\partial x}\hat{e}_x + \frac{\partial}{\partial y}\hat{e}_y + \frac{\partial}{\partial z}\hat{e}_z. \tag{1.5}$$

This definition can be extended to higher dimensions as well.

The above definition only works really on a *scalar field*, a function from the 3 dimensional space to the real numbers. However, one encounters vector fields as well, which calls for another definition. A *vector field*, in simple terms, is a function that assigns a vector to every point in space.

Definition 1.3. The *divergence* of a vector field $\vec{v}(\vec{r}) = v_x(x,y,z)\hat{e}_x + v_y(x,y,z)\hat{e_y} + v_z(x,y,z)\hat{e}_z$, where v_i is a scalar field, is defined as

$$\vec{\nabla} \cdot \vec{v} := \frac{\partial v_x}{\partial x} + \frac{\partial v_y}{\partial y} + \frac{\partial v_z}{\partial z}. \tag{1.6}$$

The operator carries over from the previous definition.

In Einstein's notation, the components can be rewritten as $\frac{\partial v_i}{\partial x_i} = \partial_i v_i$. A cross product may also be performed on vector fields, leading to the following definition.

Definition 1.4. The *curl* of a vector field $\vec{v}(\vec{r})$ is defined as

$$\vec{\nabla} \times \vec{v} := \left(\frac{\partial v_z}{\partial y} - \frac{\partial v_y}{\partial z}\right)\hat{e}_x + \left(\frac{\partial v_x}{\partial z} - \frac{\partial v_z}{\partial x}\right)\hat{e}_y + \left(\frac{\partial v_y}{\partial x} - \frac{\partial v_x}{\partial y}\right)\hat{e}_z. \tag{1.7}$$

The operator carries over from the previous definitions.

In Einstein's notation, the components can be rewritten as $[\vec{\nabla} \times \vec{v}]_i = \epsilon_{ijk}\partial_j v_k$, where j and k are cycled over.

1.1.2 The Laws: Mechanics of a Particle

We first discuss the third law of action-reaction pairs. Simply stated, every action has an equal and opposite reaction. It is important to note that the two action-reaction forces never act on the same body, and hence do not cancel each other out. If a human stands on a spring scale in an elevator, the human exerts a normal force $N_{h/s}$ on the spring, and the spring exerts a normal force $N_{s/h}$ on the human. The spring scale actually measures the force $N_{s/h}$. If the elevator were to accelrate upwards, then the force $N_{h/s}$ remains the same, yet the force $N_{s/h}$ increases, leading to a higher reading on the scale. One must also note that equal in mangnitude and opposite direction does not necessarily mean that they act along the same line. The second law of motion states $\vec{F} = m\vec{a}$, where $\vec{F}$ is the net force acting on a body, m is the mass of the body, and $\vec{a}$ is the acceleration of the body.

To formalize this, we discuss the mechanics of a single particle. The particle is at position vector $\vec{r}$, and moves with the velocity vector $\vec{v} = \frac{d\vec{r}}{dt}$. The *momentum* of the particle is defined as $\vec{p} := m\vec{v}$. The net force acting on the particle is defined as $\vec{F} := \frac{d\vec{p}}{dt}$. If the mass is constant, *i.e.*, independent of time, then we can write

$$\vec{F} = m\frac{d\vec{v}}{dt} = m\vec{a} = m\frac{d^2\vec{r}}{dt^2}. \tag{1.8}$$

We call our reference frame an *inertial frame* if $\vec{F} = \frac{d\vec{p}}{dt}$ holds true when independently measured. If the net force acting on the particle is zero, then $\frac{d\vec{p}}{dt} = 0$, implying that the momentum $\vec{p}$ is conserved. This is one of the conservation laws.

The *angular momentum* of the particle is defined as $\vec{L} := \vec{r} \times \vec{p}$. The net *torque* acting on the particle is defined as $\vec{\tau} := \vec{r} \times \vec{F} = \vec{r} \times \frac{d}{dt}(m\vec{v})$. If we again assume that the mass is constant, we then have

$$\frac{d\vec{L}}{dt} = \frac{d}{dt}(\vec{r} \times \vec{p}) = \vec{r} \times \frac{d\vec{p}}{dt} = \vec{r} \times \vec{F} = \vec{\tau}. \tag{1.9}$$

Thus, if $\vec{\tau} = 0$, then $\vec{L}$ is conserved. This is another conservation law. Note that in both the conservation laws, the mass is taken to be constant.

Now, suppose that the particle moves from position i, for initial, to position f, for final, taking $d\vec{s}$ steps, and gets acted upon by a force $\vec{F}$ for time dt. The *work* done by the force is defined as $W := \int_i^f \vec{F}\cdot d\vec{s}$. If the mass is constant, then we have

$$W = \int_i^f m\vec{a}\cdot d\vec{s} = m\int_i^f \frac{d\vec{v}}{dt}\cdot \vec{v}dt = \frac{m}{2}\int_i^f dt\frac{d}{dt}(\vec{v}\cdot\vec{v}) = \frac{m}{2}\left(v_f^2 - v_i^2\right) = T_f - T_i, \tag{1.10}$$

where T denotes the *kinetic energy* of the particle. This is, roughly stated, the *work-energy theorem*. Again, note that the mass is taken to be constant. If the exeternal work done W is zero, then T is conserved. This is our third conservation law. If the work done between two positions is independent of the path taken, then the force is termed a *conservative force*. If the path starts and ends at the same point, then the work done is zero in case of a conservative force. Thus, $\oint_{\mathcal{C}} \vec{F}\cdot d\vec{s} = 0$, where $\mathcal{C}$ is a closed path. The opposite is also true; a necessary and sufficient condition for a force to be conservative is $\vec{F} = -\vec{\nabla}V$, where V is a scalar field called the *potential energy*. In this case,

$$W = \oint_{\mathcal{C}} \vec{F}\cdot d\vec{s} = -\oint_{\mathcal{C}} \vec{\nabla}V\cdot d\vec{s} = -\oint_{\mathcal{C}} dV = 0. \tag{1.11}$$

By the work-energy theorem, for an open path, we have

$$\int_i^f \vec{F}\cdot d\vec{s} = -\int_i^f \vec{\nabla}V\cdot d\vec{s} = -(V_f - V_i) = T_f - T_i \implies T_f + V_f = T_i + V_i. \tag{1.12}$$

Thus, the quantity $T+V$ is always conserved in a conservative field; this quantity is termed the *total energy*, $E = T + V$.

Galilean Transformation

July 23rd.

Suppose one observer sits at a frame of reference O and another obsever sits at a frame of reference O', which is moving away from O with a constant velocity $\vec{v}$. In this case, the positions of the two observers can be related as $\vec{r'} = \vec{r} - \vec{v}t$. However, this is based on the assumption that $t' = t$; time is experienced the same by both the observers. Newton made this assumption without any justification and it seemed right at the time. It is still important to note that Newton's form of $\vec{F} = m\frac{d^2\vec{r}}{dt^2}$ holds true in both cases—it is independent of the frame of reference. Unless the velocity $\vec{v}$ is much, much higher (comparable to the speed of light), we can make confidently make the assumption $t = t'$ and work with Newton's equations. A transformation of coordinates between two frames of reference which differ only by constant relative motion is termed a *Galilean transformation*.

All in all, *this* is where mechanics ends of a single particle mass. Note that we have not introduced the notion of internal or external forces. For much more complicated bodies made up of many more particles, more theory has to be introduced. We now discuss a system of particles.

1.1.3 The Laws: System of Particles

The first step is to choose a frame of reference and a coordinate system, and stick to it. Every particle is then assigned a position in time $\vec{r}_i$, a velocity $\dot{\vec{r}}_i$, and its acceleration $\ddot{\vec{r}}_i$. Along with this, each particle experiences an external force $\vec{F}_i^{\text{ext}}$ and forces by the other particles $\vec{F}_{ji}$ for $j \neq i$; we make the assumption that no particle exerts a force on itself. By Newton's second law, we obtain

$$\vec{F}_i^{\text{tot}} = \vec{F}_i^{\text{ext}} + \sum_{j\neq i}\vec{F}_{ji} = \frac{d\vec{p}_i}{dt}. \tag{1.13}$$

If N_p dentes the number of particles, and we assume that $m_i(t) \equiv m_i$, summing over all particles leads to

$$\sum_{i=1}^{N_p}\sum_{j=1,j\neq i}^{N_p}\vec{F}_{ji} + \sum_{i=1}^{N_p}\vec{F}_i^{\text{ext}} = \sum_{i=1}^{N_p}\frac{d}{dt}(m_i\vec{v}_i) = \frac{d^2}{dt^2}\left(\sum_{i=1}^{N_p} m_i\vec{r}_i\right). \tag{1.14}$$

We invoke the weak form of Newton's third law, which states $\vec{F}_{ji} + \vec{F}_{ij} = 0$ for all pairs $i \neq j$. With this, it is easy to see that the first summation term of the above equations drops out. Let us denote the total mass of the system by M_{tot}. Thus, we then have

$$\implies \sum_{i=1}^{N_p} \vec{F}_i^{\text{ext}} = \frac{d^2}{dt^2}\left(\frac{\sum_{i=1}^{N_p} m_i \vec{r}_i}{\sum_{i=1}^{N_p} m_i} \cdot M_{\text{tot}}\right) \tag{1.15}$$

$$\implies \vec{F}_{\text{tot}}^{\text{ext}} = M_{\text{tot}} \cdot \frac{d^2}{dt^2}\left(\frac{\sum_{i=1}^{N_p} m_i \vec{r}_i}{\sum_{i=1}^{N_p} m_i}\right) = M_{\text{tot}} \ddot{\vec{R}}_{CM} \tag{1.16}$$

where the quantity $\vec{R}_{CM}$ is termed the *center of mass* of the system. For the total momentum of the system, we work roughly the same—

$$\vec{P}_{\text{tot}} = \sum_{i=1}^{N_p} \vec{p}_i = M_{\text{tot}} \cdot \frac{d}{dt}\left(\frac{\sum_{i=1}^{N_p} m_i \vec{r}_i}{M_{\text{tot}}}\right) = M_{\text{tot}} \dot{\vec{R}}_{CM}. \tag{1.17}$$

Differentiating the above with respect to time gives us Newton's second law of motion for a system of particles.

$$\frac{d\vec{P}_{\text{tot}}}{dt} = M_{\text{tot}} \frac{d^2 \vec{R}_{CM}}{dt^2} = \vec{F}_{\text{tot}}^{\text{ext}}. \tag{1.18}$$

Thus, $\vec{P}_{\text{tot}}$ is conserved if the total external force is zero. This is our conservation of momentum law for a system of particles. Similarly, we can derive the law of conservation of angular momentum by defining $\vec{L}_i = \vec{r}_i \times \vec{p}_i$, and $\vec{L}_{\text{tot}} = \sum_{i=1}^{N_p} \vec{L}_i$. Noting that $\frac{d\vec{r}}{dt} \times \vec{p} = \vec{v} \times m\vec{v} = 0$, we get

$$\frac{d\vec{L}_{\text{tot}}}{dt} = \frac{d}{dt}\left(\sum_{i=1}^{N_p} \vec{r}_i \times \vec{p}_i\right) = \sum_{i=1}^{N_p} \vec{r}_i \times \frac{d\vec{p}_i}{dt} = \sum_{i=1}^{N_p} \vec{r}_i \times \vec{F}_i^{\text{ext}} + \sum_{i,j=1,j\neq i}^{N_p} \vec{r}_i \times \vec{F}_{ji} \tag{1.19}$$

$$\implies \frac{d\vec{L}_{\text{tot}}}{dt} = \sum_{i=1}^{N_p} \vec{r}_i \times \vec{F}_i^{\text{ext}} = \sum_{i=1}^{N_p} = \vec{\tau}_i^{\text{ext}} = \vec{\tau}_{\text{tot}}^{\text{ext}}. \tag{1.20}$$

In the above, we have made use of the strong form of Newton's third law, which states that the opposite and equal forces also act on the same line joining them. If the total external torque vanishes, then the total angular momentum is conserved.

July 28th.

For the i^{th} particle, its position vector and the positive vector of the center of mass are related as $\vec{r}_i = \vec{R}_{CM} + \vec{r}_i'$, where $\vec{r}_i'$ is the relative position between them. Taking the time derivative gives us $\vec{v}_i = \vec{v}_{CM} + \vec{v}_i'$, giving the instantaneous relative velocity between them. Using this, we rewrite as

$$\vec{L}_{\text{tot}} = \sum_{i=1}^{N_p} m_i(\vec{r}_i' + \vec{R}_{CM}) \times (\vec{v}_i' + \vec{v}_{CM}). \tag{1.21}$$

Looking at the center of mass terms ($\vec{R}_{CM}$ and $\vec{v}_{CM}$), we have

$$T_1 = \left(\sum_{i=1}^{N_p} m_i\right) \vec{R}_{CM} \times \vec{v}_{CM} = \vec{R}_{CM} \times (M_{\text{tot}} \vec{v}_{CM}) = \vec{R}_{CM \times \vec{P}_{\text{tot}}} = \vec{L}_{CM,\text{O}}, \tag{1.22}$$

where $\vec{L}_{CM,\text{O}}$ denotes the angular momentum of the center of mass about the chosen origin. Looking at the relative terms ($\vec{r}_i'$ and $\vec{v}_i'$), we have

$$T_2 = \sum_{i=1}^{N_p} m_i \vec{r}_i' \times \vec{v}_i' = \sum_{i=1}^{N_p} \vec{r}_i' \times \vec{p}_i' = \sum_{i=1}^{N_p} \vec{L}_{i,CM}, \tag{1.23}$$

where $\vec{L}_{i,CM}$ denotes the angular momentum of the i^{th} particle about the center of mass. Looking at the terms $\vec{r}_i'$ and $\vec{v}_{CM}$ leads to an evaluation of zero. Similarly, looking at the terms $\vec{R}_{CM}$ and $\vec{v}_i'$ also results to an evaluation of zero. Thus, with all terms taken care of, we have

$$\vec{L}_{\text{tot}} = \vec{L}_{CM,\text{O}} + \sum_{i=1}^{N_p} \vec{L}_{i,CM}. \tag{1.24}$$

Recall that $W_{if} = T_f - T_i$ holds regardless of the nature of the force. If the force is conservative, then potential is introduced and $T_f + V_f = T_i + V_i$ holds true even when W_{if} is non-zero. We translate this for a system of particles.

$$W_{i\mapsto f} = \int_i^f \sum_{i=1}^{N_p} \vec{F}_i \cdot d\vec{S}_i, \tag{1.25}$$

where $\vec{F}_i = \vec{F}_i^{\text{ext}} + \sum_{j=1,j\neq i}^{N_p} \vec{F}_{ji}$ (*Note that the i in the integral bound represents the initial state and has no relation to the indexing i.*). Thus, we have

$$W_{i\mapsto f} = \int_i^f \sum_{i=1}^{N_p} \vec{F}_i^{\text{ext}} d\vec{S}_i + \int_i^f \sum_{i=1}^{N_p} \left(\sum_{j=1,i\neq j}^{N_p} \vec{F}_{ji} \cdot d\vec{S}_i \right). \tag{1.26}$$

Let us focus on the first term for now.

$$\int_i^f \sum_{i=1}^{N_p} \vec{F}_i^{\text{ext}} d\vec{S}_i = \int_i^f \sum_{i=1}^{N_p} m_i \dot{\vec{v}}_i \cdot \vec{v}_i dt = \sum_{i=1}^{N_p} \int_i^f d\left(\frac{1}{2} m_i v_i^2\right) = \sum_{i=1}^{N_p} \int_i^f dT_i. \tag{1.27}$$

If we define $T_{\text{tot}} = \sum_{i=1}^{N_p} T_i$, we get

$$W_{i\mapsto f} = T_{\text{tot},f} - T_{\text{tot},i}. \tag{1.28}$$

This equation is the work-energy theorem for a system of particles. Before dealing with the second term, let us try to alter the kinetic energies to relate to the center of mass instead. We have

$$T_{\text{tot}} = \frac{1}{2}\sum_{i=1}^{N_p} m_i(\vec{v}_{CM} + \vec{v}_i') \cdot (\vec{v}_{CM} + \vec{v}_i') = \frac{1}{2} M_{\text{tot}} \left|\vec{v}_{CM}\right|^2 + \frac{1}{2}\sum_{i=1}^{M_p} m_i \left|\vec{v}_i'\right|^2. \tag{1.29}$$

We now make our first assumption: *the external forces are conservative.* Thus,

$$\int_i^f \sum_{i=1}^{N_p} \vec{F}_i^{\text{ext}} \cdot d\vec{S}_i = -\int_i^f \vec{\nabla}_i V_i^{\text{ext}} \cdot d\vec{S}_i = \sum_{i=1}^{N_p} \left(-\int_i^f \vec{\nabla}_i V_i^{\text{ext}} \cdot d\vec{S}_i \right) \tag{1.30}$$

Our next assumption is that the force $\vec{F}_{ij}$, for all $i \neq j$, is conservative. This implies that for all pairs $i \neq j$, there exists V_{ij} such that $\vec{F}_{ij} = -\vec{\nabla}_j V_{ij}$. Another assumption we make is the fact that $\vec{F}_{ij}$ is central; that the potential V_{ij} between the i^{th} and j^{th} particles must only depend on the absolute distance between them. That is, $V_{ij} = V_{ij}(|\vec{r}_i - \vec{r}_j|) = V_{ji}(|\vec{r}_i - \vec{r}_j|)$. Thus, from both these assumptions, we get $\vec{F}_{ij} = -\vec{\nabla}_j V_{ij} \Leftrightarrow \vec{F}_{ji} = -\vec{\nabla}_i V_{ij}$. This then shows that $\vec{F}_{ji} = -\vec{F}_{ij}$, so our assumptions are valid. Also, it can be shown that the gradient can be rewritten as

$$\vec{\nabla}_j V_{ij}(|\vec{r}_i - \vec{r}_j|) = (\vec{r}_i - \vec{r}_j) f(\vec{r}_{ij}). \tag{1.31}$$

Let $\vec{r} = \vec{r}_i = \vec{r}_j$. Then, the j^{th} component is

$$\left[\vec{\nabla}_j V_{ij}(|\vec{r}|)\right]_j = \frac{\partial V_{ij}}{\partial r}\left[\vec{\nabla}_j r\right]_j = \frac{\partial}{\partial x_j}\left((x_i - x_j)^2 + (y_i - y_j)^2 + (z_i - z_j)^2\right)^{1/2} = \frac{-(x_i - x_j)}{r}. \tag{1.32}$$

Therefore,

$$\vec{\nabla}_j V_{ij}(|\vec{r}|) = \left(-\frac{1}{r}\frac{\partial V_{ij}}{\partial r}\right)((x_i - x_j)\hat{e}_x + (y_i - y_j)\hat{e}_y + (z_i - z_j)\hat{e}_z) = f(\vec{r}_{ij})(\vec{r}_i - \vec{r}_j) \tag{1.33}$$

as desired. Coming back to the second term of the work equation, we now have

$$\sum_{i,j=1;i\neq j}^{N_p} \int_i^f \vec{F}_{ji} \cdot d\vec{S}_i = \sum_{1\leq i<j\leq N_p} \int_i^f \left(F_{ij} \cdot d\vec{S}_j + \vec{F}_{ji} \cdot d\vec{S}_i\right). \tag{1.34}$$

Focusing on just the integral for now gives us

$$\int_i^f \left(F_{ij} \cdot d\vec{S}_j + \vec{F}_{ji} \cdot d\vec{S}_i\right) = -\int_i^f \left((\vec{\nabla}_j V_{ij}) \cdot d\vec{S}_j + (\vec{\nabla}_i V_{ij}) \cdot d\vec{S}_i\right) = -\int_i^f \vec{\nabla}_{ij} V_{ij} \cdot d\vec{S}_{ij} \tag{1.35}$$

where $d\vec{S}_{ij} = d\vec{S}_i - d\vec{S}_j$. Thus, the second term is finally

$$-\sum_{i,j=1;i\neq j}^{N_p} \frac{1}{2} \int_i^f \vec{\nabla}_{ij} V_{ij} \cdot d\vec{S}_{ij} \tag{1.36}$$

which gives us

$$T_{\text{tot},f} - T_{\text{tot},i} = -\sum_{i=1}^{N_p} \left(V_{i,\text{final}}^{\text{ext}} - V_{i,\text{init}}^{\text{ext}}\right) + (-\frac{1}{2}) \sum_{i,j=1;i\neq j}^{N_p} (V_{ij}^{\text{final}} - V_{ij}^{\text{init}}) = -(V_{\text{tot}}, \text{init}^{\text{sys}} - V_{\text{tot,final}}^{\text{sys}}). \tag{1.37}$$

We conclude that the total energy, $E_{\text{sys}}^{\text{tot}} = T_{\text{tot}}^{\text{sys}} + V_{\text{tot}}^{\text{sys}}$ is conserved.

1.2 The Laws: Rigid Bodies

July 30th.

We now extensively study the mechanics of rigid bodies. By a *rigid body*, we mean one where the distance between any pair of particles remains constant; for all i, j, we have

$$|\vec{r}_i(t) - \vec{r}_j(t)|^2 = c_{ij}^2. \tag{1.38}$$

Now comes the question of the coordinate system. For a rigid body, the cartesian coordinate system may not always be the nicest to work with. For example, suppose we have a particle always travelling on the surface of a sphere of radius R. Then its coordinates are always related as $x^2(t) + y^2(t) + z^2(t) = R^2$. To come up with a more suitable coordinate system, we look at 2 dimensions first.

1.2.1 Coordinate Systems

If a particle is at (x, y), then we can rewrite its coordinates as (r, ϕ), where r is the distance from the origin, and ϕ is the counter-clockwise angle made with the positive x-axis. Here, they are related as $x = r\cos\phi$ and $y = r\sin\phi$. If the particle moves around in a circle, then choosing the latter coordinate system, known as the *polar coordinate system*, proves to be useful since $\dot{r}(t) = 0$ and only the analysis on ϕ is to be done.

Extending this idea into three dimensions, we first assign a new unit vector $\hat{e}_r$ that points in the same direction as $\vec{r}$. θ is an angle measured from the positive z-axis, and takes the values $[0, \pi]$, with $\theta = 0$ at the north pole and $\theta = \pi$ at the south pole. The unit vector associated with θ is $\hat{e}_\theta$ which points tangentially along the longitude from the north to the south. Next, ϕ is termed the angle the positive x-axis makes with the projection of $\vec{r}$ onto the xy-plane in a counter-clockwise manner. Thus, $\phi \in [0, 2\pi]$. The unit vector associated with ϕ is $\vec{e}_\phi$ in such a way that $\hat{e}_r \times \hat{e}_\theta = \hat{e}_\phi$. This is the *spherical coordinate system*. It is important to note that the unit vectors are *not* fixed and are, in fact, functions of r, θ, and ϕ.

Remark 1.5. The cartesian coordinates and the spherical coordinates are related as

$$x(t) = r(t)\sin\theta(t)\cos\phi(t), \tag{1.39}$$
$$y(t) = r(t)\sin\theta(t)\sin\phi(t), \tag{1.40}$$
$$z(t) = r(t)\cos\theta(t). \tag{1.41}$$

If $r(t) = R$ is fixed, then $x^2 + y^2 + z^2 = R^2$ and, thus,

$$\dot{x}(t) = R\dot{\theta}(t)\cos\theta(t)\cos\phi(t) - R\dot{\phi}(t)\sin\theta(t)\sin\phi(t), \tag{1.42}$$

$$\dot{y}(t) = R\dot{\theta}(t)\cos\theta(t)\sin\phi(t) + R\dot{\phi}(t)\sin\theta(t)\cos\phi(t), \tag{1.43}$$

$$\dot{z}(t) = -R\dot{\theta}(t)\sin\theta(t). \tag{1.44}$$

Our next goal is to derive expressions for gradient, divergence, and curl in the coordinates (r, θ, ϕ). For any function f of (x, y, z), it can also be a function f of (r, θ, ϕ). One might be tempted to think that the gradient, or divergence or curl, can simply be found by replacing $x, y, z, \hat{e}_x, \hat{e}_y, \hat{e}_z$ with $r, \theta, \phi, \hat{e}_r, \hat{e}_\theta, \hat{e}_\phi$ in the equation(s), but then one has the problem of dimensions. Figuring out the exact equations is left as an exercise the reader.

In the *cyclindrical coordinate system*, one has (ρ, ϕ, z), where ϕ and z retain their usual meaning from the spherical and cartesian coordinates respectively, and ρ is the distance of the coordinate from the z-axis. In this case, the unit vector associated with ρ satisfies $\hat{e}_\rho \times \hat{e}_\phi = \hat{e}_z$. The reader is urged to formulate the relationships between the three coordinate systems.

1.2.2 Generalized Coordinates

Coming back to the rigid body mechanics, the inter-particle distance equation can be rewritten as

$$(x - x_C(t))^2 + (y - y_C(t))^2 + (z - z_C(t))^2 - R^2 = 0. \tag{1.45}$$

Such an equation which can be written in the form

$$f(\vec{r}_1, \vec{r}_2, \ldots, \vec{r}_{N_p}, t) = 0 \tag{1.46}$$

is termed a *holonomic constraint*. These are again of two types; a mechanical system is *rheonomous* if its equations of constraints contain time as an explicit variable. These are hard to deal with. On the other hand, a mechanical system is *scleronomous* if the equations of constraints do not contain the time as an explicit variable.

We now discuss the notion of *generalized coordinates*, starting with N_p particles in d dimensions. Thus, we will have dN_p independent coordinates to deal with. For now, we deal with $d = 3$. Suppose we also have k equations of constraints which are holonomic in nature. Combining, we than have

$$(3N_p - k) \text{ independent `generalized' coordinates} \tag{1.47}$$

which are termed $\{q_j\}$.

Equilibrium

August 4th.

The case of *equilibrium* simply means $\vec{F}_i = 0$ for $i = 1, 2, \ldots, N_p$. Trivially, the vector sum $\sum_{i=1}^{N_p} \vec{F}_i$ is also zero. More importantly, $\vec{F}_i \cdot \delta\vec{r}_i = 0$ and no work is being done by the system. Here, $\delta\vec{r}_i$ is the *virtual displacement* of each particle and are not independent of each other. The virtual displacements are consistent with the constraints of the system.

$$\sum_{i=1}^{N_p} (\vec{F}_i^{(a)} + \vec{f}_i) \cdot \delta\vec{r}_i = 0 \tag{1.48}$$

where $\vec{F}_i^{(a)}$ is the applied force, and $\vec{f}_i$ is the constraint force. We now make our first assumption: that the constraint forces do no work under virtual displacements. That is, $\vec{f}_i \cdot \delta\vec{r}_i = 0$. Of course, this implies that

$$\sum_{i=1}^{N_p} \vec{F}_i^{(a)} \cdot \delta\vec{r}_i = 0 \tag{1.49}$$

This is the *principle of virtual work*: the total virtual work done by the applied forces vanishes. Note that this does not imply that each applied force is zero. We now extend this to the case of dynamics. In this case, Newton's second law gives us $\vec{F}_i = \dot{\vec{p}}_i$, where $\vec{p}_i$ is the linear momentum of the ith particle. Then we have:

$$\vec{F}_i - \dot{\vec{p}}_i = 0 \Rightarrow \sum_{i=1}^{N_p} (\vec{F}_i^{(a)} + \vec{f}_i - \dot{\vec{p}}_i) \cdot \delta\vec{r}_i = 0 \tag{1.50}$$

Again invoking the assumption that the constraint forces do no virtual work, we find:

$$\sum_{i=1}^{N_p} (\vec{F}_i^{(a)} - \dot{\vec{p}}_i) \cdot \delta\vec{r}_i = 0 \tag{1.51}$$

This is known as *d'Alembert's principle*. It states that the difference between the applied forces and the time derivative of the momentum (i.e., the inertial forces) does no virtual work.

Suppose $\vec{r}_i = \vec{r}_i(q_1, \ldots, q_n, t)$ for all $i = 1, 2, \ldots, N_p$, with n generalized coordinates. We transfer many definitions to the idea of generalized coordinates

The first step is the definition of the velocities. We have

$$\vec{v}_i = \frac{d\vec{r}_i}{dt} = \sum_{j=1}^{n} \frac{\partial \vec{r}_i}{\partial q_j} \frac{dq_j}{dt} + \frac{\partial \vec{r}_i}{\partial t} \tag{1.52}$$

The virtual displacement is again a function of the generalized coordinates, so we can write this as

$$\delta\vec{r}_i = \sum_{j=1}^{n} \frac{\partial \vec{r}_i}{\partial q_j} \delta q_j. \tag{1.53}$$

Moving forward,

$$\sum_{i=1}^{N_p} \vec{F}_i^{(a)} \cdot \delta\vec{r}_i = \sum_{i=1}^{N_p} \vec{F}_i^{(a)} \cdot \left(\sum_{j=1}^{n} \frac{\partial \vec{r}_i}{\partial q_j} \delta q_j \right) = \sum_{j=1}^{n} \left(\sum_{i=1}^{N_p} \vec{F}_i^{(a)} \cdot \frac{\partial \vec{r}_i}{\partial q_j} \right) \delta q_j = \sum_{j=1}^{N} Q_j \delta q_j \tag{1.54}$$

where Q_j is the scalar $\sum_{i=1}^{N_p} \vec{F}_i^{(a)} \cdot \frac{\partial \vec{r}_i}{\partial q_j}$. Next, using Newtons' law,

$$\dot{\vec{p}}_i \cdot \delta\vec{r}_i = m_i \ddot{\vec{r}}_i \cdot \delta\vec{r}_i \Rightarrow \sum_{i=1}^{N_p} m_i \ddot{\vec{r}}_i \cdot \delta\vec{r}_i = \sum_{j=1}^{n} \left(\sum_{i=1}^{N_p} m_i \ddot{\vec{r}}_i \cdot \frac{\partial \vec{r}_i}{\partial q_j} \right) \delta q_j. \tag{1.55}$$

A bit of algebra shows us

$$m_i \ddot{\vec{r}}_i \cdot \frac{\partial \vec{r}_i}{\partial q_j} = \frac{d}{dt} \left(m_i \dot{\vec{r}}_i \cdot \frac{\partial \vec{r}_i}{\partial q_j} \right) - m_i \dot{\vec{r}}_i \cdot \frac{d}{dt} \left(\frac{\partial \vec{r}_i}{\partial q_j} \right) \tag{1.56}$$

and

$$\frac{d}{dt} \left(\frac{\partial \vec{r}_i}{\partial q_j} \right) = \frac{\partial}{\partial q_j} \left(\frac{d\vec{r}_i}{dt} \right) = \frac{\partial \vec{v}_i}{\partial q_j}. \tag{1.57}$$

Denoting the time derivative of q_j as $\dot{q}_j$, we have

$$\frac{\partial \vec{v}_i}{\partial \dot{q}_j} = \frac{\partial}{\partial \dot{q}_j} \left(\frac{d\vec{r}_i}{dt} \right) = \frac{\partial}{\partial \dot{q}_j} \left(\sum_{k=1}^{n} \frac{\partial \vec{r}_i}{\partial q_k} \dot{q}_k + \frac{\partial \vec{r}_i}{\partial t} \right) = \frac{\partial \vec{r}_i}{\partial q_j}. \tag{1.58}$$

Thus,

$$\frac{d}{dt} \left(m_i \ddot{\vec{r}}_i \cdot \frac{\partial \vec{r}_i}{\partial q_j} \right) = \frac{d}{dt} \left(m_i \vec{v}_i \cdot \frac{\partial \vec{v}_i}{\partial \dot{q}_j} \right) \text{ and } m_i \dot{\vec{r}}_i \cdot \frac{d}{dt} \left(\frac{\partial \vec{r}_i}{\partial q_j} \right) = m_i \vec{v}_i \cdot \frac{\partial \vec{v}_i}{\partial q_j}. \tag{1.59}$$

Note that both these terms are distinct and are summed over i and j. With $T = (\sum_{i=1}^{N_p} m_i v_i^2)/2$, we turn to

$$\sum_{i=1}^{N_p} \dot{\vec{p}}_i \cdot \delta\vec{r}_i = \sum_{i=1}^{N_p}\left(\sum_{j=1}^{N_p}\left(\frac{d}{dt}\left(m_i\vec{v}_i \cdot \frac{\partial \vec{v}_i}{\partial \dot{q}_j}\right) - m_i\vec{v}_i \cdot \frac{\partial \vec{v}_i}{\partial q_j}\right)\delta q_j\right) \tag{1.60}$$

$$= \sum_{j=1}^{n}\left(\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_j}\right) - \frac{\partial T}{\partial q_j}\right)\delta q_j. \tag{1.61}$$

This is the term that appears from the above. Remember that there was also a term $\vec{F}_i^{(a)} \cdot \delta\vec{r}_i = \sum_{j=1}^{n} Q_j \delta q_j$. Therefore,

$$\sum_{i=1}^{N_p}(\vec{F}_i^{(a)} - \dot{\vec{p}}_i)\cdot\delta\vec{r}_i = 0 \implies \sum_{j=1}^{n}\left(\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_j}\right) - \frac{\partial T}{\partial q_j} - Q_j\right)\delta q_j = 0. \tag{1.62}$$

What we have essentially done is transform the equations depending on the non-independent virtual displacements into a linear sum of δq_j's which are independent of each other. We get the desired result

$$\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_j}\right) - \frac{\partial T}{\partial q_j} - Q_j = 0 \text{ for all } j = 1, 2, \ldots, n. \tag{1.63}$$

Let us use the above result to derive the equations for three dimensional motion. Here, $q_1 = x$, $q_2 = y$, and $q_3 = z$. Also, $\frac{\partial T}{\partial q_j} = 0$ for all j, where $T = \frac{1}{2}m(\dot{x}^2 + \dot{y}^2 + \dot{z}^2)$. Moreover,

$$\frac{\partial T}{\partial \dot{q}_1} = \frac{\partial T}{\partial \dot{x}} = m\dot{x} \text{ and } \frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_1}\right) = \frac{d}{dt}(m\dot{x}) = m\ddot{x} = 0 \tag{1.64}$$

which matches the standard laws of motions. For n particles, we have

$$\frac{d}{dt}\left(\frac{\partial T}{\partial q_j}\right) - \frac{\partial T}{\partial q_j} = Q_j \text{ for all } j = 1, 2, \ldots, n. \tag{1.65}$$

Since $\vec{F}_i = -\vec{\nabla}_i V$,

$$Q_j = \sum_{i=1}^{N_p} \vec{F}_i \cdot \frac{\partial \vec{r}_i}{\partial q_j} = -\sum_{i=1}^{N_p} \vec{\nabla}_i V \cdot \frac{\partial \vec{r}_i}{\partial q_j} = -\frac{\partial V}{\partial q_j} \tag{1.66}$$

V generally does not depend on $\dot{q}_j$. So,

$$\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_j}\right) - \frac{\partial (T-V)}{\partial q_j} = 0 \implies \frac{d}{dt}\left(\frac{\partial (T-V)}{\partial \dot{q}_j}\right) - \frac{\partial (T-V)}{\partial q_j} = 0. \tag{1.67}$$

This aove equation holds true for any j. The quantity $T - V$ is termed the *Lagrangian* denoted as $L = L(\{q_j\}, \{\dot{q}_j\}, t)$.

Example 1.6. Let us look at a particle moving in 2 dimensions via polar coordinates. The force acting on it is $\vec{F} = F_x\hat{e}_x + F_y\hat{e}_y$. Here, $q_1 = r$ and $q_2 = \theta$. The kinetic energy is $\frac{1}{2}m(\dot{x}^2 + \dot{y}^2)$. The coordinates are changed to polar as $x = r\cos\theta = x(r, \theta)$ and $y = r\sin\theta = y(r, \theta)$, giving us

$$\dot{x} = \dot{r}\cos\theta - r\dot{\theta}\sin\theta \text{ and } \dot{y} = \dot{r}\sin\theta + r\dot{\theta}\cos\theta. \tag{1.68}$$

The kinetic energy then becomes

$$T = \frac{1}{2}m(\dot{x}^2 + \dot{y}^2) = \frac{1}{2}m\dot{r}^2 + \frac{1}{2}mr^2\dot{\theta}^2 \tag{1.69}$$

giving us

$$\frac{\partial T}{\partial \dot{r}} = m\dot{r} \implies \frac{d}{dt}\left(\frac{\partial T}{\partial \dot{r}}\right) = \frac{d}{dt}(m\dot{r}) = m\ddot{r}. \tag{1.70}$$

Along with $\frac{\partial U}{\partial \dot{r}} = 0$, we get

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{r}}\right) = m\ddot{r}. \tag{1.71}$$

Moreover, $\frac{\partial T}{\partial r} = mr\dot{\theta}^2$ and $\frac{\partial U}{\partial r} = -F_r$. Thus, we have

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{r}}\right) - \frac{\partial L}{\partial r} = 0 \implies m\ddot{r} - mr\dot{\theta}^2 - F_r = 0. \tag{1.72}$$

Moving on to $q_2 = \theta$, we first have

$$\frac{\partial T}{\partial \dot{\theta}} = mr^2\dot{\theta} \implies \frac{d}{dt}\left(\frac{\partial T}{\partial \dot{\theta}}\right) = 2mr\dot{r}\dot{\theta} + mr^2\ddot{\theta} = \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{\theta}}\right) \tag{1.73}$$

since $\frac{\partial U}{\partial \dot{\theta}} = 0$. Moreover, since $\frac{\partial T}{\partial \theta} = 0$, we get $\frac{\partial L}{\partial \theta} = -\frac{\partial U}{\partial \theta} = rF_\theta = r\vec{F} \cdot \hat{e}_\theta$. Thus, we get

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{\theta}}\right) - \frac{\partial L}{\partial \theta} = 0 \implies 2mr\dot{r}\dot{\theta} + mr^2\ddot{\theta} = rF_\theta. \tag{1.74}$$

The final differential equations we derived are known as the *Euler-Lagrangian equations of motion*.

1.3 Working with the Euler-Lagrangian Equations

August 6th.

For a system of N_p particles, k constraint equations, moving in d dimensions, the number of generalized coordinates is given by $dN_p - k = n$. The kinetic energy T is expressed as

$$T = \sum_{i=1}^{N_p} \frac{1}{2} m_i \left|\dot{\vec{r}}_i\right|^2. \tag{1.75}$$

The Lagrangian is then formed as $L = T - V$, where V is the potential energy. Here, V depends on the generalized coordinates and we take the case of conservative forces. Then Euler-Lagrange equation is then used as

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_i}\right) - \frac{\partial L}{\partial q_i} = 0 \qquad \text{for all } i = 1, 2, \ldots, n \tag{1.76}$$

where q_i is the i^{th} generalized coordinate. We now express the kinetic energy T is another form.

Denote $\dot{\vec{r}}_i$ by $\vec{v}_i$, which can be written as

$$\vec{v}_i = \sum_{j=1}^{n} \frac{\partial \vec{r}_i}{\partial q_j}\dot{q}_j + \frac{\partial \vec{r}_i}{\partial t}. \tag{1.77}$$

Here, $\dot{q}_j$ is the exact derivative, that is, $\frac{dq_j}{dt}$. Plugging this in the kinetic energy formula, we get

$$T = \sum_{i=1}^{n} \frac{1}{2} m_i \left(\sum_{j=1}^{n} \frac{\partial \vec{r}_i}{\partial q_j}\dot{q}_j + \frac{\partial \vec{r}_i}{\partial t}\right) \cdot \left(\sum_{k=1}^{n} \frac{\partial \vec{r}_i}{\partial q_k}\dot{q}_k + \frac{\partial \vec{r}_i}{\partial t}\right). \tag{1.78}$$

T can be written as $T = T_0 + T_1 + T_2$, where the T_i's are

$$T_0 = \sum_{i=1}^{N_p} \frac{1}{2} m_i \left(\frac{\partial \vec{r}_i}{\partial t} \cdot \frac{\partial \vec{r}_i}{\partial t} \right), \tag{1.79}$$

$$T_1 = \sum_{j=1}^{n} \dot{q}_j M_j \quad \text{where} \quad M_j = \sum_{i=1}^{N_p} m_i \frac{\partial \vec{r}_i}{\partial q_j} \cdot \frac{\partial \vec{r}_i}{\partial t}, \tag{1.80}$$

$$T_2 = \sum_{1 \le j < k \le n} M_{jk} \dot{q}_j \dot{q}_k + \sum_{j=1}^{n} M_{jj} (\dot{q}_j)^2 \quad \text{where} \quad M_{jk} = \sum_{i=1}^{N_p} m_i \frac{\partial \vec{r}_i}{\partial q_j} \cdot \frac{\partial \vec{r}_i}{\partial q_k}. \tag{1.81}$$

Writing $T = T_0 + T_1 + T_2$ gives us an advantage; if $\frac{\partial \vec{r}_i}{\partial t} = 0$, then the terms T_0 and T_1, which depend on this factor, drop out.

Example 1.7. Let us discuss *Atwood's machine*, which is simply two masses M_1 and M_2 on opposite sides of a weightless pulley. If we assume that total length of the string to be L, then M_1 hangs x length below the pulley, and M_2 hands $L - x$ length below the pulley. In this case we have

$$T = \frac{1}{2} M_1 \dot{x}^2 + \frac{1}{2} M_2 (\dot{x}^2) = \frac{1}{2} (M_1 + M_2) \dot{x}^2, \tag{1.82}$$

$$V = -M_2 g (L - x) - M_1 g x, \tag{1.83}$$

$$L = \frac{1}{2} (M_1 + M_2) \dot{x}^2 + M_1 g x + M_2 g (L - x). \tag{1.84}$$

Here, our generalized coordinate is only x. The first and second terms in the Euler-Lagrange equation are

$$\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{x}} \right) = (M_1 + M_2) \ddot{x} \quad \text{and} \quad \frac{\partial L}{\partial x} = (M_1 - M_2) g. \tag{1.85}$$

Thus, the equation of motion is

$$\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{x}} \right) - \frac{\partial L}{\partial x} = 0 \implies \ddot{x} = \left(\frac{M_1 - M_2}{M_1 + M_2} \right) g. \tag{1.86}$$

Example 1.8. Suppose a wire is infinitely long and rotates uniformly about a point P on it. Also suppose a point mass (bead) is on the wire, moving about frictionless on the wire due to the rotation. Denote the particle's position with respect to P as $(r(t), \theta(t))$. Since there is no constraint force, we have $V = 0$. Denoting $\dot{\theta} = \omega$ shows us only r is a generalized coordinate. Also noting that T is identically L, we have

$$L = T = \frac{1}{2} m \dot{r}^2 + \frac{1}{2} m r^2 \omega^2. \tag{1.87}$$

Thus,

$$\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{r}} \right) = m \ddot{r} \quad \text{and} \quad \frac{\partial L}{\partial r} = m \omega^2 r. \tag{1.88}$$

The equation of motion tells us

$$\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{r}} \right) - \frac{\partial L}{\partial r} = 0 \implies \ddot{r} = \omega^2 r \implies r(t) = r_0 e^{\omega t}. \tag{1.89}$$

Here, the angular momentum $L(t) = m r^2 \omega = m \omega r_0^2 e^{2\omega t}$ showing us that there is an external torque. This torque is given by

$$\tau(t) = \frac{dL}{dt}(t) = 2 m \omega^2 r_0^2 e^{2\omega t} \implies F_\theta(t) = \frac{\tau(t)}{r(t)} = 2 m \omega^2 r_0 e^{\omega t}. \tag{1.90}$$

Chapter 2

VARIATIONAL PRINCIPLES

2.1 Integral Principle

August 11th.

Recall the assumptions we made prior to deriving the equations of motion:

1. $\vec{F}_i$ is decomposed into $\vec{F}_i^{(a)}$ and $\vec{f}_i$.
2. We have $\sum_{i=1}^{N_p} \vec{f}_i \cdot \delta\vec{r}_i = 0$.
3. Each $\vec{r}_i$ is expressed as $\vec{r}_i(q_1, q_2, \ldots, q_n, t)$, where the δq_k's are independent.

From our assumptions, there are no longer any constraint forces. Also recall the generalized forces where

$$Q_j = \sum_{i=1}^{N_p} \vec{F}_i^{(a)} \cdot \frac{\partial \vec{r}_i}{\partial q_j}, \quad j = 1, 2, \ldots, n, \tag{2.1}$$

satisfying

$$\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_i}\right) - \frac{\partial T}{\partial q_i} - Q_i = 0, \quad i = 1, 2, \ldots, n, \tag{2.2}$$

where $T = \sum_{i=1}^{N_p} \frac{1}{2} m_i \left|\dot{\vec{r}}_i\right|^2$ is the kinetic energy. If a V can be defined such that $\vec{F}_i = -\vec{\nabla}_i V$, then $L(q, \dot{q}, t) = T - V$, and

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_j}\right) - \frac{\partial L}{\partial q_j} = 0, \quad j = 1, 2, \ldots, n, \tag{2.3}$$

where $-\frac{\partial V}{\partial q_j} = Q_j$ and V does not depend on $\dot{q}_j$.

We now discuss the *integral principle*, also known as *Hamilton's principle*. As a precursor, we have $Q_j = U(\{q_j\}, \{\dot{q}_j\}, t)$ where U is a function of the generalized coordinates, velocities, and time, and $L(\{q_j\}, \{\dot{q}_j\}, t) = T - U$. The *action* is defined as

$$S = \int_{t_i}^{t_f} L(\{q_j\}, \{\dot{q}_j\}, t) dt. \tag{2.4}$$

The principle states that the physical trajectory taken by the system between two times t_i and t_f is the one for which the action S is stationary, usually an extremum.

2.1.1 Variational Calculus

We set up a function $f(y, \frac{dy}{dx}, x)$ where we have the correspondences $f \leftrightarrow x$, $q \leftrightarrow y$, and $L \leftrightarrow f$, which is analogous to writing $L(q, \dot{q}, t)$. We also identify $\frac{dy}{dx}$ with $\dot{y}$. Then, correspondingly, we have

$$J \equiv \int_{x_1}^{x_2} f(y, \dot{y}, x) dx. \tag{2.5}$$

Setting $y(x_1) = y_1$ and $y(x_2) = y_2$, we aim to minimize J with respect to y. We ask how does one *vary* $y(x)$, a function? To rectify this issue, we introduce a new parameter α in $y(x, \alpha)$ such that $y(x, 0) = y(x)$. In other words,

$$y(x, \alpha) = y(x, 0) + \alpha\eta(x), \tag{2.6}$$

where $\eta(x)$ is an arbitrary function that vanishes at the endpoints: $\eta(x_1) = \eta(x_2) = 0$. It is safe to assume here that every possible smooth path from x_1 to x_2 can be represented in this form; the neighbouring paths are parametrized by α. Thus,

$$J[f] = \int_{x_1}^{x_2} dx \cdot f(y(x, \alpha), \dot{y}(x, \alpha), x). \tag{2.7}$$

We can now differentiate J with respect to α and set it to 0 at $\alpha = 0$.

$$\frac{dJ}{d\alpha} = \int_{x_1}^{x_2} \left(\frac{\partial f}{\partial y}\frac{\partial y}{\partial \alpha} + \frac{\partial f}{\partial \dot{y}}\frac{\partial \dot{y}}{\partial \alpha} \right) dx. \tag{2.8}$$

Using the fact that $\frac{\partial \dot{y}}{\partial \alpha} = \frac{d}{dx}\left(\frac{\partial y}{\partial \alpha}\right) = \frac{\partial^2 y}{\partial x \partial \alpha}$ and noticing the dx term, the integral screams out to use integration by parts. Thus, on the rightmost term of the integral

$$\int_{x_1}^{x_2} dx \cdot \frac{\partial f}{\partial \dot{y}}\frac{d}{dx}\left(\frac{\partial y}{\partial \alpha}\right) = \left[-\int_{x_1}^{x_2} \frac{\partial y}{\partial x}\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial \dot{y}}\right) dx \right] + \frac{\partial f}{\partial \dot{y}}\frac{\partial y}{\partial \alpha}\bigg|_{x_1}^{x_2}. \tag{2.9}$$

We plug this back in the derivative of J, taking note that $\frac{\partial y}{\partial \alpha} = \eta(x)$.

$$\frac{dJ}{d\alpha} = \int_{x_1}^{x_2} \left(\frac{\partial f}{\partial y} - \frac{d}{dx}\left(\frac{\partial f}{\partial \dot{y}}\right) \right) \left(\frac{\partial y}{\partial \alpha}\right) dx = \int_{x_1}^{x_2} dx \left(\frac{\partial f}{\partial y} - \frac{d}{dx}\left(\frac{\partial f}{\partial \dot{y}}\right) \right) \eta(x) = 0. \tag{2.10}$$

Since this works for *any* η following our initial constraints, we must have

$$\frac{d}{dx}\left(\frac{\partial f}{\partial \dot{y}}\right) - \frac{\partial f}{\partial y} = 0. \tag{2.11}$$

Example 2.1. We show that, without any external conditions, the shortest path between two points (x_1, y_1) and (x_2, y_2) is a straight line. Each path segment is characterized as $(ds)^2 = (dx)^2 + (dy)^2$. Thus,

$$J = \int_{x_1}^{x_2} ds = \int_{x_1}^{x_2} \sqrt{(dx)^2 + (dy)^2} = \int_{x_1}^{x_2} dx \sqrt{1 + \left(\frac{dy}{dx}\right)^2} = \int_{x_1}^{x_2} dx (1 + \dot{y}^2)^{1/2} \tag{2.12}$$

Identifying with our equation, we have $f = \sqrt{1 + \dot{y}^2}$, and $\frac{\partial f}{\partial \dot{y}} = \frac{\dot{y}}{\sqrt{1+\dot{y}^2}}$ and $\frac{\partial f}{\partial y} = 0$. Thus, the equation tell us

$$\frac{d}{dx}\left(\frac{\dot{y}}{\sqrt{1 + \dot{y}^2}}\right) = 0 \implies \frac{\dot{y}}{\sqrt{1 + \dot{y}^2}} = C \implies \dot{y} = \sqrt{\frac{C^2}{1 - C^2}}. \tag{2.13}$$

Example 2.2. As another example, we show that the shortest path between the north pole ($\theta_1 = 0$) and south pole ($\theta_2 = \pi$) on a sphere is a great circle. Here, the movement in spherical coordinates is

$\vec{dt} = dr\hat{e}_r + rd\theta\hat{e}_\theta + r\sin\theta d\phi\hat{e}_\phi$. For a sphere, R is constant so really we have $\vec{dr} = Rd\theta\hat{e}_\theta + R\sin\theta d\phi\hat{e}_\phi$. Thus, the integral becomes

$$J = \int_{\theta_1}^{\theta_2} \sqrt{\vec{dr}\cdot\vec{dr}} = \int_{\theta_1}^{\theta_2} R\sqrt{1+\sin^2\theta\dot{\phi}^2}d\theta. \tag{2.14}$$

Thus, we identify x with θ, $y(x)$ with $\phi(\theta)$, f with $\sqrt{1+\sin^2\theta\dot{\phi}^2}$. Plugging the partials in the equation, we get

$$\frac{R\sin^2\theta\dot{\phi}}{\sqrt{1+\sin^2\theta\dot{\phi}^2}} = C \tag{2.15}$$

This works for any θ; plugging in $\theta_1 = 0$ shows that $C = 0$. Thus, we must conclude that $\dot{\phi} \equiv 0$; the condition of a great circle. Notice that our choice of north pole and south pole could be arbitrary. Thus, the shortest distance between any two points on a sphere $\theta_1 = 0$ and θ_2 is a great circle.

August 18th.

Remark 2.3. The following may be shown via variational calculus.

1. A geodesic in $\mathbb{R}^n$ is a 'straight' line.
2. A geodesic on a sphere S^2 is a 'great circle'.
3. The minimum time path, with end points fixed, freely falling under Mg is a '*brachistochrone*' cycloid, starting under rest.

Now suppose f is a function as $f(y_1(x),\dots,y_n(x),\dot{y}_1(x),\dots,\dot{y}_n(x),x)$. With $J = \int_{x_1}^{x_2} f dx$ and $y_k(x_1), y_k(x_2)$ fixed for all k, we again parameterize by α to show small changes as

$$y_k(x,\alpha) = y_k(x,0) + \alpha\eta_k(x) \tag{2.16}$$

with $\eta_k(x_1) = \eta_k(x_2) = 0$ for all $1 \le k \le n$. Thus,

$$\delta J = \frac{\partial J}{\partial\alpha}d\alpha = \int_{x_1}^{x_2}\left(\sum_{i=1}^n\left(\frac{\partial f}{\partial y_i}\frac{\partial y_i}{\partial\alpha} + \frac{\partial f}{\partial\dot{y}_i}\frac{\partial\dot{y}_i}{\partial\alpha}\right)dx\right). \tag{2.17}$$

For all i, the second term can be written as $T_2 = \frac{\partial f}{\partial\dot{y}_i}\frac{\partial}{\partial\alpha}\left(\frac{dy_i}{dx}\right)$ giving us the second integral-sum as

$$\sum_{i=1}^n\int_{x_1}^{x_2} dx\frac{\partial f}{\partial\dot{y}_i}\frac{d}{dx}\left(\frac{\partial y_i}{\partial\alpha}\right) = \sum_{i=1}^n\left[\frac{\partial f}{\partial y_i}\frac{\partial y_i}{\partial\alpha}\bigg|_{x_1}^{x_2} - \int_{x_1}^{x_2}\frac{d}{dx}\left(\frac{\partial f}{\partial\dot{y}_i}\right)\cdot\frac{\partial y_i}{\partial\alpha}dx\right]. \tag{2.18}$$

Plugging this in, we have

$$\delta J = \frac{\partial J}{\partial\alpha}d\alpha = \int_{x_1}^{x_2} dx\sum_{i=1}^n\left(\frac{\partial f}{\partial y_i} - \frac{d}{dx}\left(\frac{\partial f}{\partial\dot{y}_i}\right)\right)\delta y_i = \sum_{i=1}^n\delta y_i\left[\int_{x_1}^{x_2} dx\left[\frac{\partial f}{\partial y_i} - \frac{d}{dx}\left(\frac{\partial f}{\partial\dot{y}_i}\right)\right]\right] = 0. \tag{2.19}$$

The δy_i's are independent, giving the inner integral as zero for all i, resulting in

$$\frac{\partial f}{\partial y_i} - \frac{d}{dx}\left(\frac{\partial f}{\partial\dot{y}_i}\right) = 0. \tag{2.20}$$

2.2 Lagrange Multipliers

Finding the independent q_i's is *hard*; for a sphere, its simply the spherical coordinates without radial dependence. For a general surface, the solution is complicated and requires studying differential geometry and algebraic geometry. We have to introduce a way to introduce the constraints and eliminate the dependent coordinates. Thus introduces the concept of *Lagrange multipliers*.

Suppose we have a nice hemispherical bowl of radius R lying flat on the ground and a particle of mass m rests atop the bowl. Without the loss of generality, we assume the particle only moves in the xz-plane. Thus, $y = 0$, $x = R\sin\theta$, $z = R\cos\theta$. By writing this, we have ensured a constraint $f(x,z) = x^2 + z^2 - R^2 = 0$. Here the kinetic energy T is given by

$$T = \frac{1}{2}m(\dot{x}^2 + \dot{y}^2 + \dot{z}^2) = \frac{1}{2}(R\dot{\theta}\cos\theta)^2 + \frac{1}{2}(-R\dot{\theta}\sin\theta)^2 = \frac{1}{2}mR^2\dot{\theta}^2. \tag{2.21}$$

The potential is

$$V = mgz = mgR\cos\theta. \tag{2.22}$$

The Lagrangian is

$$L = T - V = \frac{1}{2}mR^2\dot{\theta}^2 - mgR\cos\theta. \tag{2.23}$$

The Euler-Lagrange equation is

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{\theta}}\right) - \frac{\partial L}{\partial \theta} = mR^2\ddot{\theta} - (mgR\sin\theta) = 0 \implies \ddot{\theta} = \frac{g}{R}\sin\theta. \tag{2.24}$$

Here, our constraint was fixed throughout; if we wanted to know the angle θ at which the particle leaves the surface, the constraint then would no longer be applicable.

In actuality, the particle presses a little on the bowl compressing it a tiny bit. Van der waals forces then react and push back with a normal force. Thus the distance r of the particle from the centre of the bowl changes a little, and we model this $x = r\sin\theta$ and $z = r\cos\theta$, with

$$L = \frac{1}{2}m(\dot{r}^2 + r^2\dot{\theta}^2) - mgR\cos\theta - V(r). \tag{2.25}$$

Here we get $\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{r}}\right) = m\ddot{r}$ and $\frac{\partial L}{\partial r} = mr\dot{\theta}^2 - mg\cos\theta - \frac{dV}{dr}$. The Euler-Lagrange equations are

$$m\ddot{r} - mr\dot{\theta}^2 + mg\cos\theta + \frac{dV}{dr} = 0, \tag{2.26}$$

$$mr^2\ddot{\theta} + 2mr\dot{r}\dot{\theta} - mgr\sin\theta = 0. \tag{2.27}$$

If we now put in the constraints $r(t) = R$ with $\dot{r} = 0$ and $\ddot{r} = 0$, we get

$$mR\dot{\theta}^2 = mg\cos\theta + \left.\frac{dV}{dr}\right|_R, \qquad mR^2\ddot{\theta} = mgR\sin\theta. \tag{2.28}$$

Plugging in the constraint later has given us an additional equation as $-F_N = \left.\frac{dV}{dr}\right|_R = mR\dot{\theta}^2 - mg\cos\theta$.

August 20th.

We now discuss how exactly one includes the constraints $\{f_\alpha(q_i,t) = 0 \mid \alpha = 1,2,\ldots,k,\ k < N\}$ in Hamilton's principle. This is where the method of Lagrange multipliers comes in. Suppose we have n generalized coordinates $\{q_i\}$ with m constraint equations, for $n > m$. We can then form the augmented Lagragian as

$$S = \int_1^2 dt\left(L + \sum_{j=1}^{m}\lambda_j f_i(\{q_j\},t)\right) \implies \delta S = 0 \tag{2.29}$$

where $\{\lambda_i\}$, from $i = 1$ to m, are termed *Lagrange multipliers*. This expands as

$$\delta S = \int_1^2 dt\left(\sum_{i=1}^{n}\left(\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_i}\right) - \frac{\partial L}{\partial q_i} + \sum_{j=1}^{m}\lambda_j\frac{\partial f_j}{\partial q_i}\right)\delta q_i\right) = 0. \tag{2.30}$$

We choose the m multipliers λ_j's such that the first m coefficients vanish, and we are left with $n - m$ coordinates that are now linearly independent. With L still being $T - V$, we get

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_i}\right) - \frac{\partial L}{\partial q_i} + \sum_{j=1}^{m}\lambda_j\frac{\partial f_j}{\partial q_i} = 0, \quad i = 1,2,\ldots,n. \tag{2.31}$$

Example 2.4. Consider the prior example of a point mass atop a hemispherical bowl of radius R. Here, $L = \frac{1}{2}m(\dot{x}^2 + \dot{z}^2) - mgz$. The easier of writing this is to switch to polar coordinates, so $L = \frac{1}{2}m(\dot{r}^2 + r^2\dot{\theta}^2) - mgr\cos\theta$. The only constraint here is $f = r - R$. So the augmented Lagrangian here is

$$\tilde{L} = L + \lambda(r - R) \tag{2.32}$$

The equations can be worked out to get

$$mr\ddot{r} - mr\dot{\theta}^2 + mg\cos\theta + \lambda_1 = 0, \tag{2.33}$$

$$2mr\dot{r}\dot{\theta} + mr^2\ddot{\theta} - mgr\sin\theta = 0. \tag{2.34}$$

One can then solve and plug in $r(t) = R$.

Example 2.5. For a ball of radius R rolling without slipping, the distance x travelled across the slope is given as $R\dot{\theta} = \dot{x}$. Thus, this constraint is $R\dot{\theta} - \dot{x}$.

2.3 Potential Dependent on Generalized Velocity

September 1st.

Essentially, a redefinition of the (generalized) momentum. Let us take the example of a point mass with charge moving in an electromagnetic filed. The force experienced by this particle is given by

$$\vec{F} = m\vec{a} = q(\vec{E} + \vec{v} \times \vec{B}) \tag{2.35}$$

where q is the charge on the particle, and $\vec{E}$ and $\vec{B}$ are the electric field and magnetic field respectively. Moreover, $\vec{E} \equiv \vec{E}(\vec{r}, t)$ and $\vec{B} \equiv \vec{B}(\vec{r}, t)$. One actually rewrites the magnetic field as

$$\vec{B}(\vec{r}, t) = \vec{\nabla} \times \vec{A}(\vec{r}, t) \tag{2.36}$$

where $\vec{A}(\vec{r}, t)$ is termed the *vector potential*. The magnetic field satisfies the constraint of $\vec{\nabla} \cdot \vec{B} = 0$ (In general, one can show that the curl of a divergence vector is always zero). When $\vec{\nabla} \times \vec{E} = 0$ is valid, the the vector field $\vec{E}$ can be written as $-\vec{\nabla}\varphi$. Also, for any vector field $\vec{B}$, if $\vec{\nabla} \times \vec{v}(\vec{r}, t) = 0$, then the velocity can always be written as $\vec{v}(\vec{r}, t) = \vec{\nabla} f(\vec{r}, t)$. If the above is not valid, we instead have Faraday's law as $\vec{\nabla} \times \vec{E} = -\frac{\partial \vec{B}}{\partial t}$. In particular, it can be written as

$$\vec{E}(\vec{r}, t) = -\vec{\nabla}\varphi(\vec{r}, t) - \frac{\partial \vec{A}(\vec{r}, t)}{\partial t}. \tag{2.37}$$

In this case, the velocity-dependent potential is

$$U = q\phi(\vec{r}, t) - q\vec{A}(\vec{r}, t) \cdot \vec{v} \tag{2.38}$$

which, in turn, gives the Lagrangian (in cartesian coordinates)

$$L = T - U = \frac{1}{2}m(\dot{x}^2 + \dot{y}^2 + \dot{z}^2) - q\phi(\vec{r}, t) + q\vec{A}(\vec{r}, t) \cdot \vec{v}. \tag{2.39}$$

Here in our formal notation, $q_1 = x$, $q_2 = y$, $q_3 = z$, and the corresponding time derivatives. Thus our equation of motion, for $q_i = q_1 = x$, is

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_i}\right) - \frac{\partial L}{\partial q_i} = 0. \tag{2.40}$$

Here,

$$\frac{\partial L}{\partial \dot{x}} = m\dot{x} + qA_x \implies \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}}\right) = m\ddot{x} + q\frac{\partial A_x}{\partial t} + q\left(\frac{\partial A_x}{\partial x}\dot{x} + \frac{\partial A_x}{\partial y}\dot{y} + \frac{\partial A_x}{\partial z}\dot{z}\right). \tag{2.41}$$

One can define $\vec{\nabla} A_x = \frac{\partial A_x}{\partial x}\hat{e}_x + \frac{\partial A_x}{\partial y}\hat{e}_y + \frac{\partial A_x}{\partial z}\hat{e}_z$, and $\vec{v} = \dot{x}\hat{e}_x + \dot{y}\hat{e}_y + \dot{z}\hat{e}_z$, giving us

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}}\right) = m\ddot{x} + q\frac{\partial A_x}{\partial t} + q(\vec{v}\cdot\vec{\nabla} A_x). \tag{2.42}$$

For the second term in the Lagragian, we have

$$\frac{\partial L}{\partial x} = -q\frac{\partial \phi}{\partial x} + q\left(v_x\frac{\partial A_x}{\partial x} + v_y\frac{\partial A_y}{\partial y} + v_z\frac{\partial A_z}{\partial z}\right). \tag{2.43}$$

Plugging them into the Lagrangian equation of motion, and using the fact that $\vec{E} = -\vec{\nabla}\phi - \frac{\partial \vec{A}}{\partial t} \implies E_x = -\frac{\partial \phi}{\partial x} - \frac{\partial A_x}{\partial t}$, we get

$$m\ddot{x} + q\frac{\partial A_x}{\partial t} + q\frac{\partial \phi}{\partial x} + q\left(v_x\frac{\partial A_x}{\partial x} + v_y\frac{\partial A_x}{\partial y} + v_z\frac{\partial A_x}{\partial z}\right) - q\left(v_x\frac{\partial A_x}{\partial x} + v_y\frac{\partial A_y}{\partial x} + v_z\frac{\partial A_z}{\partial x}\right) = 0 \tag{2.44}$$

$$\implies m\ddot{x} + q(-E_x) + q\left(v_y\left(\frac{\partial A_x}{\partial y} - \frac{\partial A_y}{\partial x}\right) + v_z\left(\frac{\partial A_x}{\partial z} - \frac{\partial A_z}{\partial x}\right)\right) = 0 \tag{2.45}$$

Similarly, one can show that $B_z = (\vec{\nabla}\times\vec{A})_z = \partial_x A_y - \partial_y A_x$ and $B_y = (\vec{\nabla}\times\vec{A})_y = \partial_z A_x - \partial_x A_z$. Thus, the above transforms as

$$m\ddot{x} - qE_x + q(v_y(-B_z) + v_z B_y) = 0. \tag{2.46}$$

Using the cross product for the final term, we get

$$m\ddot{x} = \left(q\vec{E} + q(\vec{v}\times\vec{B})\right)_x \tag{2.47}$$

This is the exact force required to describe the motion of a charged particle in an electromagnetic field.

2.4 Cyclic Coordinates and Conservation Laws

Recall the energies and how we write it:

$$T = \sum_{i=1}^{N}\frac{1}{2}m_i(\dot{x}_i^2 + \dot{y}_i^2 + \dot{z}_i^2), \quad U = U(\{x_i, y_i, z_i\}). \tag{2.48}$$

Here,

$$\frac{\partial L}{\partial \dot{x}_i} = m_i\dot{x}_i = p_{x,i}, \quad \frac{\partial L}{\partial \dot{y}_j} = p_{y,j}. \tag{2.49}$$

In terms of the generalized coordinates, $L = L(\{q_1, \ldots, q_n, \dot{q}_1, \ldots, \dot{q}_n, t\})$, and $p_i = \frac{\partial L}{\partial \dot{q}_i}$. For L independent of some q_i, we have $\frac{\partial L}{\partial q_i} = 0$. This q_i is termed a *cyclic coordinates*, and the equation of motion gives

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_i}\right) = 0 \implies \frac{dp_i}{dt} = 0. \tag{2.50}$$

For a rigid body, let q_j be the coordinate(s) of its center of mass. Then, the Lagrangian equation of motion (where $V \equiv V(\{q_j\})$) is

$$\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_j}\right) + \frac{\partial V}{\partial q_j} = 0 \implies \frac{dp_j}{dt} = -\frac{\partial V}{\partial q_j} = Q_j. \tag{2.51}$$

In $Q_j = \sum_{i=1}^{N}\vec{F}_i\cdot\frac{\partial \vec{r}_i}{\partial q_j}$, one has

$$\frac{\partial \vec{r}_i}{\partial q_j} = \lim_{dq_j\to 0}\frac{\vec{r}_i(q_j + dq_j) - \vec{r}_i(q_j)}{dq_j} = \lim_{dq_j\to 0}\frac{dq_j\hat{e}_n}{dq_j} = \hat{e}_n. \tag{2.52}$$

Thus, $Q_j = \sum_{i=1}^N \vec{F}_i \cdot \hat{e}_n = \vec{F}_{\text{sys}} \cdot \hat{e}_n$. Using $T = \sum_{i=1}^N \frac{1}{2} m_i v_i^2$, we get

$$p_j = \frac{\partial T}{\partial \dot{q}_j} = \sum_{i=1}^N m_r \dot{\vec{r}}_i \cdot \hat{e}_n \implies \frac{dp_j}{dt} = \frac{d}{dt}\left(\sum_{i=1}^n m_i \vec{v}_i \cdot \hat{e}_n\right) = Q_j = \vec{F}_{\text{sys}} \cdot \hat{e}_n. \tag{2.53}$$

Thus the Newton's law statement has been translated into the generalized coordinates statement. Now suppose q_j is cyclic and dq_j is rotation about an axis. Here, we again take $\frac{\partial T}{\partial q_j} = 0$ and $\frac{\partial V}{\partial q_j} = 0$. Thus, the equation of motion gives

$$\dot{p}_j = \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_j}\right) = \frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_j}\right) = \frac{\partial L}{\partial q_j} = -\frac{\partial V}{\partial q_j} = Q_j. \tag{2.54}$$

Working as above, $Q_j = \sum_{i=1}^N \vec{F}_i \cdot \frac{\partial \vec{r}_i}{\partial q_j}$, and

$$\frac{\partial \vec{r}_i}{\partial q_j} = \lim_{dq_j \to 0} \frac{\vec{r}_i(q_j + dq_j) - \vec{r}_i(q_j)}{dq_j} = \lim_{dq_j \to 0} \frac{dq_j \hat{e}_n}{dq_j} = \hat{e}_n. \tag{2.55}$$

To proceed, one has

$$|\partial \vec{r}_i| = |\vec{r}_i| \sin\theta dq_j \implies \left|\frac{\partial \vec{r}_i}{\partial q_j}\right| = |\vec{r}_i|\,|\hat{e}_n| \sin(\theta) \implies \frac{\partial \vec{r}_i}{\partial q_j} = \hat{e}_n \times \vec{r}_i. \tag{2.56}$$

where θ is the angle between $\vec{r}_i$ and $\hat{e}_n$. This tells us $Q_j = \sum_{i=1}^N \vec{F}_i \cdot (\hat{e}_n \times \vec{r}_i) = \sum_{i=1}^N (\vec{r}_i \times \vec{F}_i) \cdot \hat{e}_n = \vec{\tau}_{\text{tot,sys}} \cdot \hat{e}_n$.

$$p_j = \frac{\partial T}{\partial \dot{q}_j} = \sum_{i=1}^N m_i \vec{v}_i \cdot (\hat{e}_n \times \vec{r}_i) = \sum_{i=1}^N \hat{e}_n \cdot \vec{L}_i = \vec{L}_{\text{tot}} \cdot \hat{e}_n \implies \dot{p}_j = \frac{d}{dt}\left(\sum_{i=1}^N \vec{L}_i \cdot \hat{e}_n\right) = Q_j = \vec{\tau}_{\text{sys}} \cdot \hat{e}_n. \tag{2.57}$$

September 3rd.

For $L(\{q_i\}, \{\dot{q}_i\}, t)$, we have

$$\frac{dL}{dt} = \sum_{i=1}^n \left(\frac{\partial L}{\partial q_i}\dot{q}_i + \frac{\partial L}{\partial \dot{q}_i}\ddot{q}_i\right) + \frac{\partial L}{\partial t} = \frac{d}{dt}\left(\sum_{i=1}^n \dot{q}_i \frac{\partial L}{\partial \dot{q}_i}\right) + \frac{\partial L}{\partial t} \implies \frac{d}{dt}\left(\sum_{i=1}^n \dot{q}_i \frac{\partial L}{\partial \dot{q}_i} - L\right) + \frac{\partial L}{\partial t} = 0. \tag{2.58}$$

Here, one then defines the *energy function* as $h(\{q_i\}, \{\dot{q}_i\}, t) = \sum_{i=1}^n \dot{q}_i \frac{\partial L}{\partial \dot{q}_i} - L = \sum_{i=1}^n \dot{q}_i p_i - L$, resulting in

$$\frac{dh}{dt} = -\frac{\partial L}{\partial t}. \tag{2.59}$$

Thus, if the Lagrangian doesn't have an explicit time dependence, the energy function h is conserved. Physically, a time-translation invariance results in the conservation of h. Recall the expression for the total kinetic energy $T = T_0 + T_1 + T_2$ where

$$T_0 = \sum_{i=1}^N \frac{1}{n} m_i \left|\frac{\partial \vec{r}_i}{\partial t}\right|^2, \quad T_1 = \sum_{j=1}^n \dot{q}_j M_j, \quad T_2 = \sum_{j=1}^n \sum_{k=1}^n \dot{q}_j \dot{q}_k M_{jk} \tag{2.60}$$

and

$$M_j = \sum_{i=1}^N m_i \frac{\partial \vec{r}_i}{\partial t} \cdot \frac{\partial \vec{r}_i}{\partial q_j}, \quad M_{jk} = \sum_{i=1}^N \frac{1}{2} m_i \frac{\partial \vec{r}_i}{\partial q_j} \cdot \frac{\partial \vec{r}_i}{\partial q_k}. \tag{2.61}$$

Now, if $\vec{r}_i = \vec{r}_i(\{q_j\})$ for all $i = 1, 2, \ldots, N$, then $\frac{\partial \vec{r}_i}{\partial t} = 0$ tells us that $T = T_2$. Let us call this our first condition (i). For our second condition (ii), it may so happen that the Lagrangian $L = L_0(\{q\}, t) + L_1(\{q\}, \{\dot{q}\}, t) + L_2(\{q\}, \{\dot{q}\}, t)$, where L_0 is independent of the generalized velocities, L_1 is linear in the generalized velocities, and L_2 is quadratic in the generalized velocities. One can show that if $f(\{x_i\})$ is homogenous in degree n,

$$\sum_{i=1}^n x_i \frac{\partial f}{\partial x_i} = nf. \tag{2.62}$$

Applying this to the Lagrangian $L = L_0 + L_1 + L_2$, we find

$$h(\{q_i\}, \{\dot{q}_i\}, t) = \sum_{i=1}^{n} \dot{q}_i \frac{\partial L}{\partial \dot{q}_i} - L = (2L_2 + L_1) - (L_0 + L_1 + L_2) = L_2 - L_0. \tag{2.63}$$

Our this third condition (iii) is that the potentail V is independent of the velocities $\{\dot{q}_k\}$. With these three conditions, one has $L_2 = T_2 = T$ and $L_0 = -V$, showing

$$h = L_2 - L_0 = T - (-V) = T + V = E, \tag{2.64}$$

the total energy. Note that $\dot{h} = -\frac{\partial L}{\partial t} = 0$ is *not* equivalent to total energy conserved; the prior statements says that the energy function is conserved. Conditions (i), (ii), and (iii) must all be satisfied for the energy function to be the total energy.

Chapter 3

CENTRAL FORCE MOTION AND ROTATIONAL PHYSICS

3.1 The Central Force Problem

3.1.1 One Body Problem

September 15th.

Consider two particles of masses lying m_1 and m_2 lying at positions $\vec{r}_1$ and $\vec{r}_2$ respectively. If $\vec{r} = \vec{r}_2 - \vec{r}_1$ is the relative position vector, and $\vec{R}$ is the position vector of the centre of mass, then the kinetic energy may be given as

$$T = \frac{1}{2}m_1\dot{\vec{r}}_1^2 + \frac{1}{2}m_2\dot{\vec{r}}_2^2 = \frac{1}{2}(m_1+m_2)\dot{\vec{R}}^2 + \frac{1}{2}\left(\frac{m_1m_2}{m_1+m_2}\right)\dot{\vec{r}}^2. \tag{3.1}$$

Here the quantity $\mu = \frac{m_1m_2}{m_1+m_2}$ is called the *reduced mass*. If the potential energy of the system is $U(\vec{r})$, only dependent on the relative position, then the Lagrangian of the system may be written as

$$L = T - U = \frac{1}{2}M\dot{R}^2 + \frac{1}{2}\mu\dot{r}^2 - U(\vec{r}) \tag{3.2}$$

Note that R is a cyclic coordinate here. So

$$\frac{\partial L}{\partial\dot{\vec{R}}} = M\dot{\vec{R}} = \vec{P}_{CM} \implies \frac{d}{dt}\vec{P}_{CM} = 0. \tag{3.3}$$

The momentum of the center of mass is conserved. For the most popular of cases, we consider $U(\vec{r}) \equiv U(r)$, i.e., the potential energy depends only on the magnitude of the relative position vector. This is called a *conservative central force*. If a particle of mass m_1 is fixed at the origin, then the other particle of mass m_2 moves in the potential $U(r)$ with reduced mass μ. We shall focus our analysis only on the reduced mass since the center of mass positive is a cyclic coordinate. The Lagrangian of the system is

$$L = \frac{1}{2}\mu(\dot{r}^2 + r^2\dot{\theta}^2 + r^2\sin^2\theta\dot{\phi}^2) - U(r). \tag{3.4}$$

Since the problem is spherically symmetric, the angular momentum $\vec{L} = \vec{r}\times\mu\dot{\vec{r}}$ is a conserved quantity. We can infer that the motion is planar in central force (first property). Choose the axis of $\vec{L}$ to be the z-axis. Then $\theta = \frac{\pi}{2}$. The Lagrangian reduces to

$$L = T - U = \frac{1}{2}\mu(\dot{r}^2 + r^2\dot{\phi}^2) - U(r). \tag{3.5}$$

Here, ϕ is a cyclic coordinate and hence the angular momentum $\ell = \mu r^2\dot{\phi}$ is conserved, giving us $\dot{\phi} = \frac{\ell}{\mu r^2}$. From this conservation of ℓ, it also follows that $\frac{d}{dt}(\frac{1}{2}\mu r^2) = 0 \implies \frac{1}{2}r^2\dot{\phi}$ is constant. We can interpret this as $\frac{1}{2}r(r\dot{\phi}) = dA$, the infinitesimal area swept out by the radius vector in time dt. The *areal velocity* $\frac{dA}{dt}$ is constant; equal area is swept out in equal time (second property). The energy function is

$$h = \frac{1}{2}\mu\dot{r}^2 + \frac{\ell^2}{2\mu r^2} + U(r) \implies \frac{dh}{dt} = \mu\ddot{r} - \frac{\ell^2}{\mu r^3} + \frac{dU}{dr} = 0. \tag{3.6}$$

Rewriting gives us

$$\mu\ddot{r} = -\frac{d}{dr}\left(U + \frac{\ell^2}{2\mu r^2}\right) = -\frac{dU_{\text{eff}}(r)}{dr} \tag{3.7}$$

where the new quantity is termed the *effective potential energy*. The quantity $T_{\text{eff}} = \frac{1}{2}\mu\dot{r}^2$ is termed the effective kinetic energy. The total energy is $h = T_{\text{eff}} + U_{\text{eff}}$. Summarising all the properties, the formal solution for the problem is

$$\ell = \mu r^2\dot{\phi},\ E = \frac{1}{2}\mu\dot{r}^2 + \frac{\ell^2}{2\mu r^2} + U,\ \dot{r} = \sqrt{\frac{2}{\mu}\left(E - U - \frac{\ell^2}{2\mu r^2}\right)}. \tag{3.8}$$

Thus, r can be found as a function of t by integrating

$$\int_{r_0}^{r} \frac{dr}{\sqrt{\frac{2}{\mu}\left(E - U - \frac{\ell^2}{2\mu r^2}\right)}} = \int_0^t dt = t \tag{3.9}$$

and inverting the solution to get a function $r(t)$. The angle ϕ can be found as a function of t by integrating

$$\frac{d\phi}{dt} = \frac{\ell}{\mu r^2} \implies \phi - \phi_0 = \frac{\ell}{\mu}\int_0^t \frac{dt}{r^2(t)}. \tag{3.10}$$

The energy-distance diagrams using $U_{\text{eff}} = \frac{\ell^2}{2\mu r^2} + U(r)$ are very useful in visualising the motion of the particle. The points where $E = U_{\text{eff}}$ are the turning points of the motion. If $E < 0$, the motion is bounded, and if $E > 0$, the motion is unbounded. The point where U_{eff} is minimum is a point of stable equilibrium. For a potential of the form $U(r) = -\frac{k}{r}$, this 'orbit' of stable equilbrium can be found by setting $\frac{dU_{\text{eff}}}{dr} = 0$ and $\frac{d^2U_{\text{eff}}}{dr^2} > 0$. This gives us $r_0 = \frac{\ell^2}{\mu k}$. But setting $E = U_{\text{eff}}(r_0)$ gives us the effective force f_{eff} to be zero, or $\ddot{r} = 0$. This means that the radius is constant, and the motion is circular.

September 17th.
With $u = \frac{1}{r}$, and $\frac{du}{d\phi} = -\frac{1}{r^2}\frac{dr}{d\phi}$, the differential equation can be written as differential operators; using this, we shall rewrite our equation of motion.

$$\frac{d}{dt} = \frac{\ell}{\mu r^2}\frac{d}{d\phi} \implies \mu\ddot{r} = \mu\frac{d}{dt}\left(\frac{dr}{dt}\right) = \frac{\ell}{r^2}\frac{d}{d\phi}\left(\frac{\ell}{\mu r^2}\frac{dr}{d\phi}\right) = -\frac{\ell^2}{\mu}u^2\frac{d^2u}{d\phi^2}. \tag{3.11}$$

The second term is simply $\frac{\ell^2}{\mu r^3} = \frac{\ell^2}{\mu}u^3$. The final term is $\frac{dU}{dr} = \frac{dU}{du}\frac{du}{dr} = -u^2\frac{dU}{du}$. Thus, the equation of motion becomes

$$\frac{d^2u}{d\phi^2} + u = -\frac{\mu}{\ell^2}\frac{dU}{du}. \tag{3.12}$$

Thus, given (u_0, ϕ_0, ℓ, E), we can solve for $u = u(\phi)$ and hence $r = r(\phi)$. Recall the relation between the differentials dt and dr. Rewriting dt in terms of $d\phi$, we have

$$d\phi = \frac{\ell dr}{\mu r^2\sqrt{\frac{2}{\mu}\left(E - U(r) - \frac{\ell^2}{2\mu r^2}\right)}}. \tag{3.13}$$

Integrating gives us

$$\phi - \phi_0 = -\int_{u_0}^{u} \frac{du}{\sqrt{\frac{2\mu E}{\ell^2} - \frac{2\mu U(u)}{\ell^2} - u^2}}. \tag{3.14}$$

3.1.2 The Kepler Problem

We are simply considering $U = -\frac{k}{r} = -ku$ for some $k > 0$. Replacing this in the above equation gives

$$\phi - \phi_0 = -\int_{u_0}^{u} \frac{du}{\sqrt{\frac{2\mu E}{\ell^2} + \frac{2\mu k}{\ell^2}u - u^2}}. \tag{3.15}$$

Recalling that

$$\int \frac{dx}{\sqrt{\alpha + \beta x + \gamma x^2}} = \frac{1}{\sqrt{-\gamma}} \cos^{-1}\left(-\frac{(\beta + 2\gamma x)}{\sqrt{\beta^2 - 4\alpha\gamma^2 a}}\right) + C, \tag{3.16}$$

we obtain

$$\phi - \phi_0 = \cos^{-1}\left(\frac{\frac{\ell^2 u}{\mu k} - 1}{\sqrt{1 + \frac{2E\ell^2}{\mu k^2}}}\right). \tag{3.17}$$

Inverting the equation gives

$$\frac{1}{r} = u = \frac{\mu k}{\ell^2}\left(1 + \sqrt{1 + \frac{2E\ell^2}{\mu k^2}} \cos(\phi - \phi_0)\right) \tag{3.18}$$

or just simply

$$\frac{1}{r} = C\left(1 + e\cos(\phi - \phi_0)\right), \quad C = \frac{\mu k}{\ell^2},\ e = \sqrt{1 + \frac{2E\ell^2}{\mu k^2}}. \tag{3.19}$$

The above is simply the equation of a general conic section with one focus at the origin. Note that if $E > 0$, then $e > 1$ resulting in the shape of a hyperbola. If $E = 0$, then $e = 1$ resulting in the shape of a parabola. If $E < 0$ such that $0 < e < 1$ then the orbit is elliptical. To get a circle, the eccentricity must be zero, which means that the energy is *fixed*, as $E = -\frac{\mu k^2}{2\ell^2}$. Coming in with this energy results in a circular orbit.

3.2 Rotational Motion

September 24th.

In two dimensional motion, rotational motion can simply be thought of as the group action of the unitary group of matrices $U(1)$ on $\mathbb{R}^2$. This is an abelian group, and is simple to work with. In three dimensions, however, one must consider the group action of special orthogonal matrices $SO(3)$ on $\mathbb{R}^3$. This is a non-abelian group, and hence the order of rotations matters.

For rotation on a plane, the direction of $\vec{L}$ is fixed, and we simply have

$$\vec{L} = \left|\vec{L}\right| \hat{e}_L, \quad \frac{d\vec{L}}{dt} = \frac{d\left|\vec{L}\right|}{dt}\hat{e}_L. \tag{3.20}$$

Note that since $\vec{L} = \vec{r} \times \vec{p}$, the angular momentum depends on the choice of the origin. Taking the time derivative of the total angular momentum of a system of particles gives

$$\frac{d\vec{L}_{\text{tot}}}{dt} = \frac{d}{dt}\sum_{i=1}^{N} \vec{r}_i \times \vec{p}_i = \sum_{i=1}^{N}\left(\frac{d\vec{r}_i}{dt} \times \vec{p}_i + \vec{r}_i \times \frac{d\vec{p}_i}{dt}\right) = \sum_{i=1}^{N} \vec{r}_i \times \vec{F}_i = \vec{\tau}_{\text{tot}}. \tag{3.21}$$

Let us consider rotation only about a fixes axis, say about $\hat{e}_z$. The angular velocity may be denoted by $\vec{\omega} = \omega\hat{e}_z$. The velocity of a particle at position $\vec{r}_i$ is given by $\vec{v}_i = \vec{\omega} \times \vec{r}_i$. In the simplest case, there is only one particle of mass m at position $\vec{r}$ and velocity $\vec{v}$; its angular momentum will simply be $\vec{L} = \vec{r} \times m\vec{v}$. For a continuous body S, with $\vec{p} = (dm)\vec{v}$, the angular momentum is given by

$$\vec{L} = \int_S \vec{r} \times \vec{p} = \int_S (\vec{r} \times \vec{v})dm = \int_S (\vec{r} \times (\vec{\omega} \times \vec{r}))dm. \tag{3.22}$$

Recalling the identity $\vec{a}\times(\vec{b}\times\vec{c}) = (\vec{a}\cdot\vec{c})\vec{b} - \vec{c}(\vec{a}\cdot\vec{b})$, we obtain

$$\vec{L} = \int_S (\vec{r}\cdot\vec{r})\vec{\omega}dm = \left(\int_S r^2 dm\right)\omega\hat{e}_z = I_{zz}\omega\hat{e}_z \tag{3.23}$$

where the quantity $I_{zz} = \int_S r^2 dm$ is termed the *moment of inertia* about the z-axis. The subscript zz is used is lieu of future generalizations. Thus, the relation gives $L = I\omega$ for rotation about a fixes axis. For a discrete case, $I_{zz} = \sum_{i=1}^N m_i r_i^2$. The kinetic energy of the rotating body, using $dT = \frac{1}{2}v^2(dm)$, is given by

$$T = \int_S dT = \int_S \frac{1}{2}(\omega r)^2 dm = \frac{\omega^2}{2}\int_S r^2 dm = \frac{1}{2}I_{zz}\omega^2. \tag{3.24}$$

The continuous version of the center of mass formula is

$$\left(\frac{\sum_i m_i\vec{r}_i}{\sum_i m_i} =\right)\ \vec{r}_{CM} = \frac{\int_S \vec{r}dm}{\int_S dm}. \tag{3.25}$$

We, earlier, obtained the equations $\vec{r} = \vec{r}_{CM} + \vec{r}'$ and $\vec{v} = \vec{v}_{CM} + \vec{v}'$, where the primed quantities are relative to the center of mass. Working as we did before in chapter 1, and using $\vec{p}_{CM} = M\vec{v}_{CM}$ and $\vec{v}' = \vec{\omega}'\times\vec{r}'$, we obtain

$$\vec{L} = \vec{r}_{CM}\times\vec{p}_{CM} + \left(\int_S r'^2 dm\right)\omega'\hat{e}_z \ \text{ or } \ \vec{L} = \vec{L}_{CM} + I_{CM}^{zz}\omega'\hat{e}_Z. \tag{3.26}$$

The kinetic energy, when the center of mass transates and the rotations is $\vec{\omega}' = \omega'\hat{e}_z$, is given by

$$T = \frac{1}{2}Mv_{CM}^2 + \frac{1}{2}I_{CM}^{zz}\omega'^2. \tag{3.27}$$

If we consider rotation about an origin at another point O with axis parallel to the rotation axis through the center of mass, then the kinetic energy gives

$$T = \frac{1}{2}Mr_{CM}^2\omega^2 + \frac{1}{2}I_{CM}^{zz}\omega^2 = \frac{1}{2}(I_{CM}^{zz} + Mr_{CM}^2)\omega^2 = \frac{1}{2}I_O^{zz}\omega^2. \tag{3.28}$$

This is called the *parallel axis theorem*, giving $I_O^{zz} = I_{CM}^{zz} + Mr_{CM}^2$. The torque about O is given by

$$\vec{L}_O = \sum_{i=1}^N m_i(\vec{r}_i - \vec{r}_O)\times(\dot{\vec{r}}_i - \dot{\vec{r}}_O) \implies \frac{d\vec{L}_O}{dt} = \sum_{i=1}^N(\vec{r}_i - \vec{r}_O)\times(\vec{F}_i^{\text{ext}} + \vec{F}_i^{\text{int}}). \tag{3.29}$$

One can show that the internal force vanishes for forces acting along $\vec{r}_i - \vec{r}_j$. Thus, we have

$$\frac{d\vec{L}_O}{dt} = \vec{\tau}_{\text{tot},O}^{\text{ext}} = \sum_{i=1}^N m_i(\vec{r}_i - \vec{r}_O)\times(-\ddot{\vec{r}}_O) = \vec{\tau}_{\text{tot},O}^{\text{ext}} + (\vec{r}_{CM} - \vec{r}_O)\times(-M\ddot{\vec{r}}_O). \tag{3.30}$$

This last term, $-M\ddot{\vec{r}}_O$, is called the *fictitious force* or the *pseudoforce*.

3.2.1 Rotation in Three Dimensions

September 29th.

We will restrict ourselves to rigid bodies, where the distance between any two particles is a constant. We start by discussing the *Mozzi-Chasles theorem*, which states that the most general motion of a rigid body can be represented as a rotation about an axis and a translation along the same axis. So, if we consider a point P of the body, the entire motion can be inferred from the translation of P and a single rotation about an axis through P. Thus the required quantities are $\vec{r}_P$, $\vec{v}_P$, and $\vec{\omega}$. For any other point Q of the body, we have

$$\vec{r}_Q = \vec{r}_P + \vec{r}_{Q/P}, \quad \dot{\vec{r}}_Q = \dot{\vec{r}}_P + \dot{\vec{r}}_{Q/P}, \quad \ddot{\vec{r}}_Q = \ddot{\vec{r}}_P + \ddot{\vec{r}}_{Q/P}. \tag{3.31}$$

If P is fixed and we perform an infinitesimal motion of any other point with no turning back, the *Borsuk-Ulam theorem* states that there must be two antipodal points which do not move at all under this infinitesimal motion.

The quantities above are related as $\vec{v}_p = \vec{\omega} \times \vec{r}_P(t)$, with $|\vec{v}_P| = |\vec{\omega}|\,|\vec{r}_P| \sin\theta$, where θ is the angle between $\vec{\omega}$ and $\vec{r}_P$. Suppose we have three bodies S_1, S_2, and S_3, with the relative angular velocity of S_1 with respect to S_2 being $\vec{\omega}_{12}$, and so on. Then the angular velocity of S_1 with respect to S_3 is given by

$$\vec{\omega}_{13} = \vec{\omega}_{12} + \vec{\omega}_{23}. \tag{3.32}$$

Now suppose we have a rigid body S rotating about an axis, with $\vec{\omega} = \omega_1 \hat{e}_x + \omega_2 \hat{e}_y + \omega_3 \hat{e}_z$. For an infinitesimal mass element of a point on the body, $dm = \rho(x, y, z) dxdydz$, the angular momentum of the entire body is given as

$$\vec{L} = \int_S \vec{r} \times \vec{v} dm = \int_S dm(\vec{r} \times (\vec{\omega} \times \vec{r})) = \int_S dm \left(r^2\vec{\omega} - (\vec{r} \cdot \vec{\omega})\vec{r}\right). \tag{3.33}$$

We now introduce the *inertia tensor* $\overleftrightarrow{I}$ as the linear transformation such that $\vec{L} = \overleftrightarrow{I}\vec{\omega}$. We have

$$\vec{L} = \int_S dm \left((x^2 + y^2 + z^2)(\omega_1\hat{e}_x + \omega_2\hat{e}_y + \omega_3\hat{e}_z) - (x\omega_1 + y\omega_2 + z\omega_3)(x\hat{e}_x + y\hat{e}_y + z\hat{e}_z)\right). \tag{3.34}$$

The x component of $\vec{L}$ is

$$L_x = \vec{L} \cdot \hat{e}_x = \int_S dm \left((y^2 + z^2)\omega_1 - yx\omega_2 - xz\omega_3\right). \tag{3.35}$$

Similarly, we obtain L_y and L_z. Rewriting this in matrix form, we see that

$$\begin{pmatrix} L_x \\ L_y \\ L_z \end{pmatrix} = \begin{pmatrix} \int dm(y^2 + z^2) & -\int dmxy & -\int dmzx \\ -\int dmxy & \int dm(z^2 + x^2) & -\int dmyz \\ -\int dmzx & -\int dmyz & \int dm(x^2 + y^2) \end{pmatrix} \begin{pmatrix} \omega_1 \\ \omega_2 \\ \omega_3 \end{pmatrix} \tag{3.36}$$

where the matrix thus formed is called the inertia tensor $\overleftrightarrow{I}$. The matrix here is symmetric, and is more often written as

$$I = \begin{pmatrix} I_{xx} & I_{xy} & I_{xz} \\ I_{yx} & I_{yy} & I_{yz} \\ I_{zx} & I_{zy} & I_{zz} \end{pmatrix} \tag{3.37}$$

where $I_{\mu\nu} = I_{\nu\mu}$.

Example 3.1. Consider a uniform cube of side a and mass M. The density of the cube is $\rho = \frac{M}{a^3}$. The moment of inertia about the x-axis is given by

$$I_{xx} = \int_S (y^2 + z^2) dm = \int_0^a \int_0^a \int_0^a (y^2 + z^2)\rho dxdydz = \frac{2}{3}Ma^2. \tag{3.38}$$

By symmetry, we have $I_{yy} = I_{zz} = \frac{2}{3}Ma^2$. The off-diagonal terms are given by

$$I_{xy} = I_{yx} = -\int_S xydm = -\int_0^a \int_0^a \int_0^a xy\rho dxdydz = -\frac{1}{4}Ma^2. \tag{3.39}$$

By symmetry, we have $I_{xz} = I_{zx} = I_{yz} = I_{zy} = -\frac{1}{4}Ma^2$. Thus, the inertia tensor is given by

$$\overleftrightarrow{I} = Ma^2 \begin{pmatrix} 2/3 & -1/4 & -1/4 \\ -1/4 & 2/3 & -1/4 \\ -1/4 & -1/4 & 2/3 \end{pmatrix}. \tag{3.40}$$

Let us look at the x-component of the moment of inertia. From the tensor relation above, we obtain

$$L_x = \frac{2}{3}Ma^2\omega_1 - \frac{1}{4}ML^2\omega_2 - \frac{1}{4}Ma^2\omega_3. \tag{3.41}$$

Note that even if we set $\omega_1 = \omega_2 = 0$, that is, the cube only rotates about the z-axis, we still have $L_x = -\frac{1}{4}Ma^2\omega_3 \neq 0$. Thus, $\vec{L}$ and $\vec{\omega}$ are not parallel in general.

We now move to further generalized motion. Here, the best idea is to choose P to be the centre of mass. In such a case, we know that the angular momentum is

$$\vec{L} = M(\vec{r}_{CM} \times \vec{v}_{CM}) + \int_S dm(\vec{r}' \times (\vec{\omega}' \times \vec{r}')) \tag{3.42}$$

For any arbitrary point Q, note that $\vec{\omega}_{CM/O} = \vec{\omega}' = \vec{\omega}_{Q/CM}$. Thus we have

$$M\vec{r}_{CM} \times (\vec{\omega}' \times \vec{r}_{CM}) = M\left((r_{CM}^2)\vec{\omega}' - (\vec{r}_{CM} \cdot \vec{\omega}')\vec{r}_{CM}\right) \tag{3.43}$$

We can then work as before to get relations such as $I_{xx}^{CM/O} = M(y_{CM}^2 + z_{CM}^2)$ and so on. We then obtain the *generalized parallel axis theorem* as

$$\vec{L} = (\overset{\leftrightarrow}{I}_{CM/O} + \overset{\leftrightarrow}{I}_{S/CM})\vec{\omega}' = \overset{\leftrightarrow}{I}_{S/O}\vec{\omega}'. \tag{3.44}$$

3.3 Rigid Body Rotation

October 8th.

When we start with a rigid body comprising of N particles, we are dealing with $3N$ degrees of freedom. However, a rigid body is defined by constraints of the form $|\vec{r}_i - \vec{r}_j| = r_{ij} = c_{ij}$ and still these constraints are not independent. The general idea is to choose three non-collinear points P_1, P_2, P_3 in the rigid body, which gives us 9 degrees of freedom, and the distances between then give 3 constraints, giving the final number as 6 independent degrees of freedom at any given moment.

Thus first establish position of P_1 via $x_{P_1}, y_{P_1}, z_{P_1}$ as a function of time, and then specify the position of P_2 via θ, ϕ as functions of time on the sphere centred at P_1 with radius $r = c_1$. Finally, specify the position of P_3 via λ as a function of time on the intersecting curve. Of course, this is just a general idea, and we usually use any 6 independent coordinates which are the most convenient to the problem at hand.

Suppose we have a rigid body S with which we have associated a coordinate system O' and x', y', z' which is fixed with respect to the body. We wish to relate this coordinate system with the standard one (an inertial frame) and derive our 6 generalized coordinates as so. We will location the position of $O'(t)$ and then the orientation of $(x', y', z')(t)$ with respect to (x, y, z).

If we denote $\cos\theta_{ij} = \hat{e}_{x_i'} \cdot \hat{e}_{x_j}$, (these are functions of time and are not fixed), we obtain

$$\hat{e}_{x'} = (\hat{e}_{x'} \cdot \hat{e}_x)\hat{e}_x + (\hat{e}_{x'} \cdot \hat{e}_y)\hat{e}_y + (\hat{e}_{x'} \cdot \hat{e}_z)\hat{e}_z = \cos\theta_{11}\hat{e}_x + \cos\theta_{12}\hat{e}_y + \cos\theta_{13}\hat{e}_z, \tag{3.45}$$

$$\hat{e}_{y'} = \cos\theta_{21}\hat{e}_x + \cos\theta_{22}\hat{e}_y + \cos\theta_{23}\hat{e}_z, \tag{3.46}$$

$$\hat{e}_{z'} = \cos\theta_{31}\hat{e}_x + \cos\theta_{32}\hat{e}_y + \cos\theta_{33}\hat{e}_z. \tag{3.47}$$

If we write $\vec{r} = x\hat{e}_x + y\hat{e}_y + z\hat{e}_z = x'\hat{e}_{x'} + y'\hat{e}_{y'} + z'\hat{e}_{z'}$, utilising the above relations gives us

$$x' = x\cos\theta_{11} + y\cos\theta_{12} + z\cos\theta_{13}, \tag{3.48}$$

$$y' = x\cos\theta_{21} + y\cos\theta_{22} + z\cos\theta_{23}, \tag{3.49}$$

$$z' = x\cos\theta_{31} + y\cos\theta_{32} + z\cos\theta_{33}. \tag{3.50}$$

Using the fact that $\hat{e}_{x_i} \cdot \hat{e}_{x_j} = \delta_{ij}$, we obtain the relation

$$\sum_{l=1}^{3} \cos\theta_{lm}\cos\theta_{ln} = \delta_{mn}. \tag{3.51}$$

Note that we have written $x_i' = \sum_j a_{ij}x_j$, where $a_{ij} = \cos\theta_{ij}$. It is crucial to note that what we have written down is a linear combination. The relation $\sum_i a_{ij}a_{ik} = \delta_{jk}$ implies that the matrix $A = (a_{ij})_{3\times 3}$ is an orthogonal matrix, i.e., $A^TA = AA^T = I_3$. Let us consider rotation in a plane; in such a case, we have $\hat{e}_{x'} = \cos\phi\hat{e}_x + \sin\phi\hat{e}_y$, $\hat{e}_{y'} = -\sin\phi\hat{e}_x + \cos\phi\hat{e}_y$, and $\hat{e}_{z'} = \hat{e}_z$. The matrix A, in the case of rotation in the xy-plane, is given by

$$A_{xy} = \begin{bmatrix} \cos\phi & \sin\phi & 0 \\ -\sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{3.52}$$

About the other orthogonal axes, we find A to be of the form

$$A_{yz} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\phi & \sin\phi \\ 0 & -\sin\phi & \cos\phi \end{bmatrix}, \quad A_{zx} = \begin{bmatrix} \cos\phi & 0 & -\sin\phi \\ 0 & 1 & 0 \\ \sin\phi & 0 & \cos\phi \end{bmatrix}. \tag{3.53}$$

Of course, one may verify that if A and B are orthogonal, then so is AB. Moreover, $A^{-1} = A^T$ and the identity matrix I is also orthogonal. Thus, the set of all orthogonal matrices (of order n) forms a group $O(n)$ under matrix multiplication. Note that $\det(A) = \pm 1$. The orthogonal matrices with determinant -1 do not represent pure rotations, but rather a combination of a rotation and a reflection. The set of orthogonal matrices with determinant $+1$ form a subgroup termed the special orthogonal group $SO(n)$, and we deal with $SO(3)$ for three dimensional rotations. A special property of $SO(3)$ is that any $A \in SO(3)$ can be built up as a tranformation from infinitesimal rotations. This is what makes $SO(3)$ something known as a Lie group.

3.3.1 Euler Angles

We now deal with the standard convention of the choice of 3 angles to build up a rotation. In this convention, we use 3 subsequent rotations to build up a general (final) rotation.

October 13th.

Theorem 3.2 (*Euler's theorem of rigid body motion*)**.** *If a rigid body moves with one point fixed, the motion is a rotation.*

Our goal, then becomes, is given $A(t)$, we wish to find $\hat{n}$ and Φ such that $A(t)$ is a rotation about $\hat{n}$ by an angle Φ. We start with the fact that $A \in SO(3)$, and hence $A^T A = I$ and $\det(A) = 1$. Note that if $\vec{R}$ is the rotation axis, then $A\vec{R} = \vec{R}$ tells us that $\vec{R}$ is an eigenvector of A with eigenvalue 1. For three eigenvalues $\lambda_1, \lambda_2, \lambda_3$, we have $\lambda_1\lambda_2\lambda_3 = \det(A) = 1$. Also,

$$(A - I)A^T = I - A^T \implies \det(A - I) = \det(I - A^T) = \det(I - A) \implies \det(A - I) = 0. \tag{3.54}$$

Thus, one of the eigenvalues is 1, and such a $\vec{R}$ exists. The other two eigenvalues λ_1 and λ_2 are multiplicative inverses of each other. Note that A is real; if λ is an eigenvalue, then so is λ^*. Thus, the other two eigenvalues are complex conjugates of each other, and hence of the form $e^{i\Phi}$ and $e^{-i\Phi}$.

In the case of all eigenvalues begin unity, the rotation is simply the identity. In the case of $\lambda_1 = \lambda_2 = -1 = e^{-i\Phi}$, we have $\Phi = \pi$, and the rotation is a rotation by π about some axis. In the general case where $A_{ij} = a_{ij}$, and $\vec{R} = (x, y, z)^t$, we have

$$\begin{pmatrix} a_{11} - 1 & a_{12} & a_{13} \\ a_{21} & a_{22} - 1 & a_{23} \\ a_{31} & a_{32} & a_{33} - 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = 0. \tag{3.55}$$

This gives us two independent equations, and we can add the constraint $x^2 + y^2 + z^2 = 1$ to obtain $\vec{R}$. There exists a similarity transformation $A' = X^{-1}AX$ such that

$$A' = \begin{pmatrix} \cos\Phi & \sin\Phi & 0 \\ -\sin\Phi & \cos\Phi & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{3.56}$$

The advantage of a similarity transformation is that it preserves eigenvalues, and hence the trace. Thus, $\operatorname{tr} A = 1 + 2\cos\Phi$, giving

$$\Phi = \cos^{-1}\left(\frac{1}{2}(\operatorname{tr} A - 1)\right), \quad 0 \le \Phi \le 2\pi. \tag{3.57}$$

Thus, given A, we can find Φ. Finding $\vec{R}$, however, is not as straightforward. The matrix equation above subject to the unit vector constraint gives us $\vec{R}$. There is, of course, an ambiguity in the sense that $-\vec{R}$ and $-\Phi$ are also valid solutions.

There exist two representations of A, one in terms of $\hat{e}_n$ and Φ, and the other in terms of Euler angles R_ψ, R_θ, R_ϕ. For $\vec{r} \mapsto \vec{r}' = A\vec{r}$ we wish to write $\vec{r}'$ in terms of $\vec{r}, \hat{e}_n$, and Φ. To this end, let $\vec{r}' = \vec{OQ}$ and

$\vec{R} = \vec{OP}$. Let N be the point on the rotation axis such that $ON \perp NP, NQ$, and let V be the projection of Q onto the line NP. Then

$$\vec{r}' = \vec{ON} + \vec{NV} + \vec{VQ}. \tag{3.58}$$

By definition of N, $\vec{ON} = (\vec{r} \cdot \hat{e}_n)\hat{e}_n$. Also,

$$\left|\vec{NV}\right| = \left|\vec{NQ}\right| \cos\Phi = \left|\vec{NP}\right| \cos\Phi = |\vec{r} - (\vec{r}\cdot\hat{e}_n)\hat{e}_n| \cos\Phi \implies \vec{NV} = (\vec{r} - \hat{e}_n(\hat{e}_n \cdot \vec{r}))\cos\Phi. \tag{3.59}$$

Moreover, $\left|\vec{VQ}\right| = \left|\vec{NQ}\right| \sin\Phi = \left|\vec{NP}\right| \sin\Phi = |\vec{r}| \sin\psi \sin\Phi$, where ψ is the angle between $\vec{r}$ and $\hat{e}_n$. But this is just $|\vec{r} \times \hat{e}_n| \sin\Phi$, and hence

$$\vec{VQ} = (\vec{r} \times \hat{e}_n)\sin\Phi. \tag{3.60}$$

Thus, we substitute into $\vec{r}'$ to get

$$\vec{r}' = (\vec{r}\cdot\hat{e}_n)\hat{e}_n + (\vec{r} - (\vec{r}\cdot\hat{e}_n)\hat{e}_n)\cos\Phi + (\vec{r}\times\hat{e}_n)\sin\Phi = \vec{r}\cos\Phi + (\vec{r}\cdot\hat{e}_n)\hat{e}_n(1-\cos\Phi) + (\vec{r}\times\hat{e}_n)\sin\Phi. \tag{3.61}$$

Note that the vector $\vec{r}$ isn't special, and can be replaced by any vector undergoing the same rotational transformation for a finite angle. It is left as an exercise to the reader to find a relationship between $(\hat{e}_n, \Phi)$ and the Euler angle rotations $(R_\psi, R_\theta, R_\phi)$. Thus a 'rotation formula' has been obtained.

3.3.2 Composition of Rotations

For an infinitesimal rotation, we have $x_i' = x_i + \sum_{j\neq i} \epsilon_{ij} x_j$, where (ϵ_{ij}) is an infinitesimal matrix (a tensor). Writing in vector form,

$$\vec{r}' = \vec{r} + \overset{\leftrightarrow}{\epsilon}\vec{r}. \tag{3.62}$$

We will drop the $\overset{\leftrightarrow}{\epsilon}$ and use ϵ to denote the infinitesimal rotation tensor. For two such infinitesimal rotations, we have

$$(I+\epsilon_1)(I+\epsilon_2) = I + \epsilon_1 + \epsilon_2 + \epsilon_1\epsilon_2, \quad (I+\epsilon_2)(I+\epsilon_1) = I + \epsilon_1 + \epsilon_2 + \epsilon_2\epsilon_1. \tag{3.63}$$

Since a rotation is orthogonal, we have $(I+\epsilon)^T = (I+\epsilon)^{-1}$. Thus, up to first order, we have

$$(I+\epsilon)^{-1} = I - \epsilon, \quad (I+\epsilon)^T = I + \epsilon^T \implies \epsilon^T = -\epsilon. \tag{3.64}$$

Thus, ϵ is an antisymmetric tensor, and hence has only 3 independent components. We can write

$$\epsilon = \begin{pmatrix} 0 & d\Omega_3 & -d\Omega_2 \\ -d\Omega_3 & 0 & d\Omega_1 \\ d\Omega_2 & -d\Omega_1 & 0 \end{pmatrix}. \tag{3.65}$$

The infinitesimal change in position is

$$d\vec{r} = \vec{r}' - \vec{r} = \epsilon\vec{r} = \begin{pmatrix} 0 & d\Omega_3 & -d\Omega_2 \\ -d\Omega_3 & 0 & d\Omega_1 \\ d\Omega_2 & -d\Omega_1 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}. \tag{3.66}$$

Thus we obtain $dx_1 = x_2(d\Omega_3) - x_3(d\Omega_2)$. Similarly, we obtain dx_2 and dx_3. Note that this is just the cross product of $\vec{r}$ with $\vec{d\Omega} = (d\Omega_1, d\Omega_2, d\Omega_3)^T$ (note that $\vec{d\Omega} \neq d\vec{\Omega}$; it is not the infinitesimal change in a vector, but rather just notation). Thus, we have

$$d\vec{r} = \vec{r} \times \vec{d\Omega}, \quad \text{or} \quad \vec{r}' = \vec{r} + \vec{r} \times \vec{d\Omega}. \tag{3.67}$$

This formula is only true for infinitesimal rotations. Let us verify this using the rotation formula. Recall

$$\vec{r}' = \vec{r}\cos\Phi + \hat{e}_n(\hat{e}_n \cdot \vec{r})(1-\cos\Phi) + (\vec{r}\times\hat{e}_n)\sin\Phi. \tag{3.68}$$

For an infinitesimal rotation, we have $\Phi = d\Phi$, $\cos d\Phi = 1$ and $\sin d\Phi = d\Phi$. Thus, we have

$$\vec{r}' = \vec{r} + (\vec{r}\times\hat{e}_n)d\Phi = \vec{r} + \vec{r}\times\vec{d\Omega}, \tag{3.69}$$

where we have taken $\vec{d\Omega} = \hat{e}_n d\Phi$. Thus, the two formulae agree. From this relation if we write $\hat{e}_n = (n_1, n_2, n_3)^T$, we obtain

$$\epsilon = \begin{pmatrix} 0 & n_3 & -n_2 \\ -n_3 & 0 & n_1 \\ n_2 & -n_1 & 0 \end{pmatrix} d\Phi. \tag{3.70}$$

3.4 Rate of Change

October 22nd.

COnsider the infinitesimal change $d\vec{G}$ of any vector $\vec{G}$ under an infinitesimal rotation $d\vec{\Omega}$. We have

$$(d\vec{G})_{\text{space}} = (d\vec{G})_{\text{body}} + d\vec{\Omega} \times \vec{G}. \tag{3.71}$$

Taking the time derivative, we obtain

$$\frac{d\vec{G}}{dt}\bigg|_{\text{space}} = \frac{d\vec{G}}{dt}\bigg|_{\text{body}} + \vec{\omega} \times \vec{G} \tag{3.72}$$

where $\vec{\omega}$ is the instantaneous angular velocity vector. This is a sort of hand-wavy derivation, but it can be made more rigorous. Suppose the body-set is rotated with respect to the space-set by an orthogonal matrix $A(t)$. Let $[A(t)]_{ij} = a_{ij}$. Then without the loss of generality, let at time t ($G'_i = G_i$), we have

$$G_i = \sum_j (A^{-1})_{ij} G'_j = \sum_j a_{ji} G'_j \implies dG_i = \sum_j a_{ji} dG'_j + \sum_j G'_j da_{ji} = dG'_i + \sum_j G'_j da_{ji}. \tag{3.73}$$

With $-\epsilon = \begin{pmatrix} 0 & -d\Omega_3 & d\Omega_2 \\ d\Omega_3 & 0 & -d\Omega_1 \\ -d\Omega_2 & d\Omega_1 & 0 \end{pmatrix}$, we have

$$-\epsilon_{ij} = -\epsilon_{ijk} d\Omega_k = \epsilon_{ikj} d\Omega_k \implies da_{ji} = \epsilon_{ikj} d\Omega_k \tag{3.74}$$

since $da_{ji} = (\epsilon^T)_{ij} = (-\epsilon)_{ij}$. Note that ϵ_{ij} denotes the components of the antisymmetric tensor, while ϵ_{ijk} is the Levi-Civita symbol. Thus,

$$dG_i = dG'_i + \sum_{j,k} G'_j \epsilon_{ikj} d\Omega_k = dG'_i + (d\vec{\Omega} \times \vec{G}')_i. \tag{3.75}$$

Note that the vector $\vec{G}$ did not really play a role; any vector would have worked. Taking the time derivative, we obtain the desired result. For example, putting in the position vector $\vec{r}$, we have

$$\vec{v}_s = \vec{v}_b + \vec{\omega} \times \vec{r}. \tag{3.76}$$

We take it a step further to derive the accelration—

$$\vec{a}_s = \frac{d\vec{v}_s}{dt}\bigg|_s = \frac{d\vec{v}_s}{dt}\bigg|_b + \vec{\omega} \times \vec{v}_s = \frac{d\vec{v}_b}{dt}\bigg|_b + \vec{\omega} \times (\vec{v}_b + \vec{\omega} \times \vec{r}) = \vec{a}_b + 2\vec{\omega} \times \vec{v}_b + \vec{\omega} \times (\vec{\omega} \times \vec{r}). \tag{3.77}$$

Define $\vec{F}_s = m\vec{a}_s$ and $\vec{F}_b = m\vec{a}_b$ to get

$$\vec{F}_b = \vec{F}_s - m(\vec{\omega} \times (\vec{\omega} \times \vec{r})) - 2m(\vec{\omega} \times \vec{v}_b). \tag{3.78}$$

The first of these addition terms is called the *centrifugal force*, while the second is called the *Coriolis force*. These are the only two pseudoforces that arise in a rotating frame with constant angular velocity.

Recall the angular momentum equation, $\vec{L} = \overleftrightarrow{I}\vec{\omega} \implies \sum_\beta I_{\alpha\beta}\omega_\beta = L_\alpha$. The rotational kinetic energy takes the form $T = \frac{1}{2}\vec{\omega} \cdot (\overleftrightarrow{I}\vec{\omega}) = \sum_{\alpha,\beta} \frac{1}{2} I_{\alpha\beta}\omega_\alpha\omega_\beta$. But note that the inertia tensor is symmetric, and hence can be diagonalized. Thus, there exists a coordinate system in which $\overleftrightarrow{I}$ is diagonal, i.e., $I_{\alpha\beta} = 0$ for $\alpha \neq \beta$. The corresponding axes are called the *principal axes of inertia*. Diagonalizing $\overleftrightarrow{I}$ gives

$$S^{-1}\overleftrightarrow{I}S = \begin{pmatrix} I_1 & 0 & 0 \\ 0 & I_2 & 0 \\ 0 & 0 & I_3 \end{pmatrix} \tag{3.79}$$

where I_i's are called the *principal moments of inertia*. Corresponding to these principal axes are unit vectors $\hat{e}_1, \hat{e}_2, \hat{e}_3$. In this coordinate system, we simply have $L_1 = I_1\omega_1$, and so on. Note that the principal axes are orthogonal, since S is an orthogonal matrix. Here we have taken $\vec{\omega} = \omega_1\hat{e}_1 + \omega_2\hat{e}_2 + \omega_3\hat{e}_3$ and $\vec{L} = L_1\hat{e}_1 + L_2\hat{e}_2 + L_3\hat{e}_3$. Moving forward, if $\vec{L}$ plays the role of $\vec{G}$, we get

$$\vec{\tau}_{\text{ext}} = \frac{d\vec{L}}{dt}\bigg|_s = \frac{d\vec{L}}{dt}\bigg|_b + \vec{\omega} \times \vec{L} \implies \frac{dL_i}{dt}\bigg|_b + \epsilon_{ijk}\omega_j L_k = \tau_{\text{ext},i}. \tag{3.80}$$

With $L_i = I_i\omega_i$, we have

$$\tau_{\text{ext},i} = I_i\frac{d\omega_i}{dt} + \epsilon_{ijk}(I_k\omega_k) \implies \tau_{\text{ext},1} = I_1\dot{\omega}_1 + \omega_2\omega_3(I_3 - I_2) \quad \text{and cyclic permutations.} \tag{3.81}$$

Index