Classical Mechanics Oral Exam Study Guide

Elijan Mastnak

2019-2020 Summer Semester

About These Notes

These notes give the answers to typical questions[1] from the oral exam required to pass the course *Klasična Mehanika* (Classical Mechanics), a required course for second-year physics students at the Faculty of Math and Physics in Ljubljana, Slovenia. I wrote the notes when studying for the exam and am making them publicly available in the hope that they might help others learning the same material. Although the exact oral exam questions are specific to the the physics program at the University of Ljubljana, the content is fairly standard for an undergraduate classical mechanics course and might be useful to others learning similar material. The most up-to-date version can be found on GitHub.

Navigation: For easier document navigation, the table of contents is "clickable", meaning you can jump directly to a section by clicking the colored section names in the table of contents. Unfortunately, the *clickable links do not work in most online or mobile PDF viewers*; you have to download the file first.

On Authorship: The material herein is far from original—it is primarily a combination of Professor David Tong's lectures on Classical Dynamics, Professor Anton Ramšak's lectures at the University of Ljubljana, and Goldstein's canonical *Classical Mechanics*. I take credit only for compiling the common exam material in one place and for typesetting the notes.

Disclaimer: Mistakes—both trivial typos and legitimate errors—are likely. Keep in mind that these are the notes of an undergraduate student in the process of learning the material himself—take what you read with a grain of salt. If you find mistakes and feel like telling me, by Github pull request, email or some other means, I'll be happy to hear from you, even for the most trivial of errors.

How to use these notes

- The sections follow the order in which the course material was taught.

- Each typical exam question is answered in a dedicated subsection, where the question is written in italics at the beginning of the subsection.

- Each question subsection may potentially be further divided into subsubsections to better organize the answer material.

---

[1]The oral exam questions are taken from the course website, included in the GitHub directory containing these notes, and at the time of writing are publically available at this link.

Contents

1 Newtonian Mechanics

1.1 Newtonian Dynamics for a Single of Particle

Summarize the Newtonian dynamics of a single particle.

1.1.1 Momentum and Force

A particle is described by its position vector $\boldsymbol{r}$ and the derivatives $\dot{\boldsymbol{r}} = \boldsymbol{v}$ and $\ddot{\boldsymbol{r}} = \boldsymbol{a}$

- A particle's *linear momentum* $\boldsymbol{p}$ is defined

$$\boldsymbol{p} = m\dot{\boldsymbol{r}}$$

  Newton's second law relates the net force $\boldsymbol{F}$ on a particle to its linear momentum $\boldsymbol{p}$ and governs the dynamics of particle motion. The general form is

$$\boldsymbol{F}(\boldsymbol{r}, \dot{\boldsymbol{r}}) = \dot{\boldsymbol{p}}$$

  For a particle with constant mass, $\boldsymbol{F} = m\ddot{\boldsymbol{r}}$. If the net force $\boldsymbol{F}$ on a particle is zero, then the particle's linear momentum $\boldsymbol{p}$ is conserved.

$$\boldsymbol{F} = 0 \implies \boldsymbol{p} = \text{constant}$$

- *Newton's third law* states $\boldsymbol{F}_{12} = -\boldsymbol{F}_{21}$: when two isolated particles interact, the forces they exude on each other are equal in magnitude and opposite in direction.

- Newton's laws only hold in *inertial frames*: an inertial frame is unaccelerated. Equivalently, an inertial frame is a frame in which a free particle with constant mass travels in straight line $\boldsymbol{r} = \boldsymbol{r}_0 + \boldsymbol{v}t$.

  In Euclidean space, there are 10 linearly independent transformations that preserve inertial frames. These are three rotations, three translations, three boosts and one time translation, and together they form the Galilean group.

  Because they are time derivatives of position, both linear momentum and force are invariant under space translations and are thus independent of the choice of origin.

- *Central forces* are a special class of forces between particles for which

  - The force between the particles acts along the line connecting the to particles, i.e. the force $\boldsymbol{F}_{12}$ is parallel to the relative position $\boldsymbol{r}_1 - \boldsymbol{r}_2$.
  - The force depends only on the distance $r = |\boldsymbol{r}_1 - \boldsymbol{r}_2|$ between the particles i.e. $\boldsymbol{F} = \boldsymbol{F}(r)$.

1.1.2 Angular Momentum and Torque

- A particle's *angular momentum* $\boldsymbol{L}$ with respect to the origin is defined as

$$\boldsymbol{L} = \boldsymbol{r} \times \boldsymbol{p} = m(\boldsymbol{r} \times \boldsymbol{v})$$

  The *rotational analog of Newton's second law* relates the torque $\boldsymbol{\tau}$ on the particle to angular momentum $\boldsymbol{L}$ via

$$\boldsymbol{\tau} = \dot{\boldsymbol{L}}$$

  If the net torque $\boldsymbol{\tau}$ on a particle is zero, then the particle's angular momentum $\boldsymbol{L}$ is conserved.

$$\boldsymbol{\tau} = 0 \implies \boldsymbol{L} = \text{constant}$$

- Torque is related to force by $\boldsymbol{\tau} = \boldsymbol{r} \times \boldsymbol{F}$, obtained via

$$\boldsymbol{\tau} = \frac{\mathrm{d}\boldsymbol{L}}{\mathrm{d}t} = \frac{\mathrm{d}}{\mathrm{d}t}[\boldsymbol{r} \times \boldsymbol{p}] = \dot{\boldsymbol{r}} \times \boldsymbol{p} + \boldsymbol{r} \times \dot{\boldsymbol{p}} = \boldsymbol{0} + \boldsymbol{r} \times \boldsymbol{F} = \boldsymbol{r} \times \boldsymbol{F}$$

since $\dot{\boldsymbol{r}} \times \boldsymbol{p} = m(\boldsymbol{v} \times \boldsymbol{v}) = \boldsymbol{0}$.

- Unlike linear momentum and force, both angular momentum and torque are measured with respect to a chosen reference point and thus depend on the choice of origin.

1.1.3   Work, Energy and Conservative Forces

- The *work* W done by a force $\boldsymbol{F}$ acting on a particle traveling along the curve γ is given by the line integral

$$W = \int_{\gamma} \boldsymbol{F} \cdot \mathrm{d}\boldsymbol{r}$$

- The *kinetic energy* of a particle of mass m moving with velocity $\boldsymbol{v}$ is

$$T = \frac{1}{2}m|\dot{\boldsymbol{r}}|^2 = \frac{mv^2}{2}$$

- The change in a particle's kinetic energy under the influence of a force $\boldsymbol{F}$ during some physical process equals the work W done by the force during that process. If the particle travels from position $\boldsymbol{r}_1$ at time t_1 to position $\boldsymbol{r}_2$ at time t_2, the change in the particle's kinetic energy is

$$T(t_2) - T(t_1) = \int_{t_1}^{t_2} \frac{\mathrm{d}T}{\mathrm{d}t}\,\mathrm{d}t = \int_{t_1}^{t_2} \boldsymbol{F} \cdot \dot{\boldsymbol{r}}\,\mathrm{d}t = \int_{\boldsymbol{r}_1}^{\boldsymbol{r}_2} \boldsymbol{F} \cdot \mathrm{d}\boldsymbol{r} = W$$

- *Conservative forces* have two special properties:

  1. They depend only on position $\boldsymbol{r}$, not on velocity $\dot{\boldsymbol{r}}$, i.e. $\boldsymbol{F} = \boldsymbol{F}(\boldsymbol{r})$
  2. The work done by the force on a particle between two points is independent of the path taken between the points

  Conservative forces can be written as the gradient of a potential energy $V(\boldsymbol{r})$. In this case, $\boldsymbol{F} = -\nabla V$. The work done by a conservative force on a particle along a curve γ equals the change in potential energy between the endpoints.

$$W = \int_{\gamma} \boldsymbol{F} \cdot \mathrm{d}\boldsymbol{r} = -(V_2 - V_1) \qquad \text{and} \qquad \oint_{\kappa} \boldsymbol{F} \cdot \mathrm{d}\boldsymbol{r} = 0$$

  where κ is a closed curve.

- For particles under the influence of conservative forces, we have the *energy conservation law*

$$T_1 + V_1 = T_2 + V_2 \equiv E$$

  stating that the particle's total energy E is conserved in any physical process between two states e.g. 1 and 2. The law arises from

$$W = T_2 - T_1 = -\int_{\boldsymbol{r}_1}^{\boldsymbol{r}_2} V(r) \cdot \mathrm{d}\boldsymbol{r} = -V_2 + V_1$$

1.2 Newtonian Dynamics for Systems of Particles

Summarize the Newtonian dynamics for a system of N particles with masses $m_1, m_2, \ldots, m_N$ and position vectors $\boldsymbol{r}_1, \boldsymbol{r}_2, \ldots, \boldsymbol{r}_N$.

1.2.1 Center of Mass

- The total mass of the system is $M = m_1 + m_2 + \cdots + m_N$. The system's *center of mass*, which we'll denote by $\boldsymbol{r}^*$, is

$$\boldsymbol{r}^* = \frac{m_1 \boldsymbol{r}_1 + m_2 \boldsymbol{r}_2 + \cdots + m_N \boldsymbol{r}_N}{m_1 + m_2 + \cdots + m_N} \implies M\boldsymbol{r}^* = \sum_i m_i \boldsymbol{r}_i$$

- The velocity and acceleration of the center of mass are, respectively

$$\boldsymbol{v}^* = \frac{\mathrm{d}\boldsymbol{r}^*}{\mathrm{d}t} \qquad \text{and} \qquad \boldsymbol{a}^* = \frac{\mathrm{d}\boldsymbol{v}^*}{\mathrm{d}t} = \frac{\mathrm{d}^2\boldsymbol{r}^*}{\mathrm{d}t^2}$$

1.2.2 Momentum and Forces

- The system's *total momentum* $\boldsymbol{p}_{\text{tot}}$ is the vector sum of the momenta $\boldsymbol{p}_i$ of the constituent particles

$$\boldsymbol{p}_{\text{tot}} = \sum_i \boldsymbol{p}_i$$

- Newton's second law still holds individually for each particle:

$$\boldsymbol{F}_i = \dot{\boldsymbol{p}}_i$$

where $\boldsymbol{F}_i$ is the net force on the ith particle. The force $\boldsymbol{F}_i$ is conventionally decomposed into the net external force and the inter-particle forces

$$\boldsymbol{F}_i = \boldsymbol{F}_i^{\text{ext}} + \sum_{j \neq i} \boldsymbol{F}_{ij}$$

where $\boldsymbol{F}_{ij}$ is the force of the jth particle on the ith particle.

- The *total force* $\boldsymbol{F}_{\text{tot}}$ on the system is

$$\boldsymbol{F}_{\text{tot}} = \sum_i \boldsymbol{F}_i = \sum_i \boldsymbol{F}_i^{\text{ext}} + \sum_{i,j;\, j \neq i} \boldsymbol{F}_{ij} = \boldsymbol{F}_{\text{ext}} + \sum_{i < j}(\boldsymbol{F}_{ij} + \boldsymbol{F}_{ji}) = \boldsymbol{F}_{\text{ext}}$$

where the term $(\boldsymbol{F}_{ij} + \boldsymbol{F}_{ji})$ vanishes by Newton's third law. The result tells us that the total force on a system equals the net external force.

- *Newton's second law* for the entire system is

$$\boldsymbol{F}_{\text{ext}} = \sum_i \boldsymbol{F}_i = \sum_i \dot{\boldsymbol{p}}_i = \dot{\boldsymbol{p}}_{\text{tot}}$$

If the net external force on the system is zero, the system's total momentum is conserved.

$$\boldsymbol{F}_{\text{ext}} = 0 \implies \boldsymbol{p}_{\text{tot}} = \boldsymbol{p}^* = \text{constant}$$

In terms of the acceleration of the center of mass, Newton's second law is

$$\boldsymbol{F}_{\text{ext}} = \sum_i \dot{\boldsymbol{p}}_i = M \sum \frac{m_i \ddot{\boldsymbol{r}}_i}{m_i} = M\boldsymbol{a}^*$$

1.2.3   Angular Momentum and Torque

- The system's *total angular momentum* $\boldsymbol{L}_{\text{tot}}$ is the vector sum of the angular momenta $\boldsymbol{L}_i$ of the constituent particles

$$\boldsymbol{L}_{\text{tot}} = \sum_i \boldsymbol{L}_i$$

The *total torque* $\boldsymbol{\tau}_{\text{tot}}$ on the system is

$$\boldsymbol{\tau}_{\text{tot}} = \sum_i \boldsymbol{r}_i \times \boldsymbol{F}_i$$

- The time derivative of the total angular momentum is

$$\dot{\boldsymbol{L}}_{\text{tot}} = \sum_i \dot{\boldsymbol{L}}_i = \sum_i [\dot{\boldsymbol{r}}_i \times \boldsymbol{p}_i + \boldsymbol{r}_i \times \dot{\boldsymbol{p}}_i] = \boldsymbol{0} + \sum_i \boldsymbol{r}_i \times \boldsymbol{F}_i$$

$$= \sum_i \boldsymbol{r}_i \times \left(\boldsymbol{F}_i^{\text{ext}} + \sum_{j \neq i} \boldsymbol{F}_{ij} \right) = \sum_i \boldsymbol{r}_i \times \boldsymbol{F}_i^{\text{ext}} + \sum_{i,j;\, j \neq i} \boldsymbol{r}_i \times \boldsymbol{F}_{ij}$$

$$= \boldsymbol{\tau}_{\text{ext}} + \sum_{i<j} (\boldsymbol{r}_i - \boldsymbol{r}_j) \times \boldsymbol{F}_{ij}$$

If the internal forces $\boldsymbol{F}_{ij}$ are central forces then $(\boldsymbol{r}_i - \boldsymbol{r}_j) \parallel \boldsymbol{F}_{ij}$ and the second term vanishes, leaving

$$\dot{\boldsymbol{L}}_{tot} = \boldsymbol{\tau}_{\text{ext}}$$

1.2.4   Work and Energy

- The *total kinetic energy* of a system of N particles is

$$T = \frac{1}{2} \sum_i m_i \dot{\boldsymbol{r}}_i^2$$

We can decompose the position vector $\boldsymbol{r}_i$ into

$$\boldsymbol{r}_i = \boldsymbol{R} + \tilde{\boldsymbol{r}}_i$$

where $\boldsymbol{R}$ is the position of the system's center of mass and $\tilde{\boldsymbol{r}}_i$ is the distance from the ith particle to the center of mass. With the decomposition $\boldsymbol{r}_i = \boldsymbol{R} + \tilde{\boldsymbol{r}}_i$, the total kinetic energy becomes

$$T = \frac{1}{2} \sum_i m_i (\dot{\boldsymbol{R}} + \dot{\tilde{\boldsymbol{r}}}_i)^2 = \frac{1}{2} \sum_i m_i \dot{\boldsymbol{R}}^2 + \frac{1}{2} \sum_i m_i \dot{\tilde{\boldsymbol{r}}}_i^2 + \dot{\boldsymbol{R}} \cdot \frac{\mathrm{d}}{\mathrm{d}t} \sum_i m_i \tilde{\boldsymbol{r}}_i$$

$$= \frac{1}{2} M \dot{\boldsymbol{R}}^2 + \frac{1}{2} \sum_i m_i \dot{\tilde{\boldsymbol{r}}}_i^2$$

where the last term vanishes because $\sum_i m_i \tilde{\boldsymbol{r}} = 0$. The result shows that the total kinetic interpretation decomposes into the kinetic energy of the center of mass and the *internal kinetic energy* describing how the system's particles move around the center of mass.

- The *total work* done on the system during a physical process is defined as

$$W = \sum_i \int \boldsymbol{F}_i \cdot \mathrm{d}\boldsymbol{r}_i$$

7

The change in the system's total kinetic energy for a physical process running between the times t_1 and t_2 equals the total work W done on the system during that process.

$$T(t_2) - T(t_1) = \int_{t_1}^{t_2} \frac{\mathrm{d}T}{\mathrm{d}t} \, \mathrm{d}t = \int_{t_1}^{t_2} \sum_i m_i \dot{\boldsymbol{r}}_i \cdot \ddot{\boldsymbol{r}}_i \, \mathrm{d}t = \sum_i \int_{t_1}^{t_2} \boldsymbol{F}_i \cdot \dot{\boldsymbol{r}}_i \, \mathrm{d}t$$

$$= \sum_i \int \boldsymbol{F}_i \cdot \mathrm{d}\boldsymbol{r}_i = W$$

- If both the external and internal forces are conservative we can write them in terms of potentials as

$$\boldsymbol{F}_i^{\mathrm{ext}} = -\nabla_i V_i(\boldsymbol{r}_1, \ldots, \boldsymbol{r}_N) \qquad \text{and} \qquad \boldsymbol{F}_{ij} = -\nabla_i V_{ij}(\boldsymbol{r}_1, \ldots, \boldsymbol{r}_N)$$

where $\nabla_i \equiv \frac{\partial}{\partial \boldsymbol{r}_i}$. In this case the change in kinetic energy is

$$T(t_2) - T(t_1) = \sum_i \int \boldsymbol{F}_i \cdot \mathrm{d}\boldsymbol{r}_i = \sum_i \left[\int \boldsymbol{F}_i^{\mathrm{ext}} \cdot \mathrm{d}\boldsymbol{r}_i + \sum_{j \neq i} \int \boldsymbol{F}_{ij} \cdot \mathrm{d}\boldsymbol{r}_i \right]$$

$$= \sum_i \left[V_i^{(1)} - V_i^{(2)} + V_{ij}^{(1)} - V_{ij}^{(2)} \right] = V^{(1)} - V^{(2)}$$

In other words, the system's total energy is conserved in the presence of both conservative internal and external forces

$$T^{(1)} + T^{(2)} = V^{(1)} + V^{(2)} \equiv E$$

1.2.5   Virial Theorem

- Consider a stable system of N particles and some quantity G defined as

$$G = \sum_i^N \boldsymbol{p}_i \cdot \boldsymbol{r}_i$$

Using the product rule, the rate of change of G with respect to time is

$$\frac{\mathrm{d}G}{\mathrm{d}t} = \sum_i \dot{\boldsymbol{p}}_i \cdot \boldsymbol{r}_i + \sum_i \boldsymbol{p} \cdot \dot{\boldsymbol{r}}_i = \sum_i \boldsymbol{F}_i \cdot \boldsymbol{r}_i + 2T$$

where $T = \frac{1}{2} \sum_i \boldsymbol{p} \cdot \dot{\boldsymbol{r}} = \frac{1}{2} \sum_i m \dot{\boldsymbol{r}}^2$ is the system's total kinetic energy.

- As long as the system is stable, it will stay in a bounded region of phase space, so G can only take on a bounded set of values. Because the value of G is bounded, the average value of its *rate of change* $\frac{\mathrm{d}G}{\mathrm{d}t}$ over any arbitrarily long time period is zero.

- Taking the time average of the equation for $\frac{\mathrm{d}G}{\mathrm{d}t}$ and applying $\left\langle \frac{\mathrm{d}G}{\mathrm{d}t} \right\rangle = 0$ gives

$$\left\langle \frac{\mathrm{d}G}{\mathrm{d}t} \right\rangle = \left\langle \sum_i \boldsymbol{F}_i \cdot \boldsymbol{r}_i \right\rangle + 2 \langle T \rangle = 0 \implies 2 \langle T \rangle = -\left\langle \sum_i \boldsymbol{F}_i \cdot \boldsymbol{r}_i \right\rangle$$

This is the *virial theorem* for a stable system of N particles.

- For interactions governed by a central potential $V = V(r)$ of the form $V(r) = kr^n$ the force can be written

$$\boldsymbol{F} = -\nabla V(r) = -\nabla(kr^n) = -knr^{n-1}\hat{\mathbf{r}}$$

The virial theorem then reads

$$2\langle T\rangle = -\left\langle \sum_i \boldsymbol{F}_i \cdot \boldsymbol{r}_i \right\rangle = -\left\langle \sum_i \left(-knr_i^{n-2}\boldsymbol{r}_i \cdot \boldsymbol{r}_i\right) \right\rangle = +\left\langle \sum_i nV(r_i) \right\rangle$$
$$= n\langle V\rangle$$

1.3   Non-Inertial Systems and Fictitious Forces

Describe the necessary modifications for Newton's laws to hold in non-inertial systems.

1.3.1   Velocity and Acceleration in a Rotating Frame

- We will be dealing with:

  – An inertial coordinate system S with coordinate axes (x, y, z)
  – An rotating coordinate system S' with coordinate axes (x', y', z')
  – A vector quantity in space (typically the position vector $\boldsymbol{r}$), which we will describe in terms of both the inertial or non-inertial coordinates.

  We assume the origins of S and S' coincide. S is at rest and S' rotates with respect to S about an axis through the mutual origin with angular velocity $\boldsymbol{\omega}$.

- From the perspective of the fixed frame S, a particle rotating with S' will appear to move with the velocity

$$\dot{\boldsymbol{r}} = \boldsymbol{\omega} \times \boldsymbol{r}$$

The key step is extending this description of rotation to the axes of S' themselves. Let $\boldsymbol{e}_1', \boldsymbol{e}_2'$ and $\boldsymbol{e}_3'$ be the unit vectors pointing along the x', y' and z' axes of S'. Just like a particle rotating with S' appeared to move with velocity $\dot{\boldsymbol{r}} = \boldsymbol{\omega} \times \boldsymbol{r}$ in S, the unit vectors rotate with velocity

$$\dot{\boldsymbol{e}}_i' = \boldsymbol{\omega} \times \boldsymbol{e}_i'$$

from the perspective of the inertial frame S.

- Let $\boldsymbol{e}_1, \boldsymbol{e}_2$ and $\boldsymbol{e}_3$ be the unit vectors pointing along the x, y and z axes of S. In the $\{\boldsymbol{e}_i\}$ basis the position of the particle in S is

$$\boldsymbol{r} = \sum_{i=1}^{3} r_i \boldsymbol{e}_i \equiv r_i \boldsymbol{e}_i$$

where the second equality adopts the shorthand summation convention. Alternatively, in the $\{\boldsymbol{e}_i'\}$ basis, the position measured in S' is

$$\boldsymbol{r} = r_i' \boldsymbol{e}_i'$$

Note that the position vector $\boldsymbol{r}$ is the same in both frames since the origins coincide, but the coordinates r_i and r_i' will in general differ, since they correspond to different bases.

- Equipped with the position vector $\boldsymbol{r}$, we can now compute the velocity with product rule for derivatives. Measured in frame S, the velocity is

$$\dot{\boldsymbol{r}} = \dot{r}_i \boldsymbol{e}_i + r_i \dot{\boldsymbol{e}}_i = \dot{r}_i \boldsymbol{e}_i$$

We have the simple expression $\dot{\boldsymbol{r}} = \dot{r}_i \boldsymbol{e}_i$ because the axes $\boldsymbol{e}_i$ do not change with time (i.e. $\dot{\boldsymbol{e}}_i = 0$) since S is inertial.

In S', using the expression $\dot{\boldsymbol{e}}_i' = \boldsymbol{\omega} \times \boldsymbol{e}_i'$, the velocity is

$$\dot{\boldsymbol{r}} = \dot{r}_i' \boldsymbol{e}_i' + r_i' \dot{\boldsymbol{e}}_i' = \dot{r}_i' \boldsymbol{e}_i' + r_i'(\boldsymbol{\omega} \times \boldsymbol{e}_i') = \dot{r}_i' \boldsymbol{e}_i' + \boldsymbol{\omega} \times (r_i' \boldsymbol{e}_i')$$
$$= \dot{r}_i' \boldsymbol{e}_i' + \boldsymbol{\omega} \times \boldsymbol{r}$$

- We now introduce a new notation to highlight the physics and hopefully clarify what is going on in the transitions between frames. We write the velocity of the particle seen by an observer in S as

$$\left(\frac{\mathrm{d}\boldsymbol{r}}{\mathrm{d}t} \right)_S = \dot{r}_i \boldsymbol{e}_i$$

Similarly, we write the velocity as seen by an observer rotating with the frame S' as

$$\left(\frac{\mathrm{d}\boldsymbol{r}}{\mathrm{d}t} \right)_{S'} = \dot{r}_i' \boldsymbol{e}_i'$$

- From the equation $\dot{\boldsymbol{r}} = \dot{r}_i' \boldsymbol{e}_i' + \boldsymbol{\omega} \times \boldsymbol{r}$, we see that observers in the two frames measure two different velocities related by

$$\left(\frac{\mathrm{d}\boldsymbol{r}}{\mathrm{d}t} \right)_S = \left(\frac{\mathrm{d}\boldsymbol{r}}{\mathrm{d}t} \right)_{S'} + \boldsymbol{\omega} \times \boldsymbol{r}$$

The discrepancy is the relative velocity $\boldsymbol{v}_{\mathrm{rel}} = \boldsymbol{\omega} \times \boldsymbol{r}$ between the two frames.

- To get the particle's acceleration we differentiate $\dot{\boldsymbol{r}}$. In S this is simply

$$\ddot{\boldsymbol{r}} = \ddot{r}_i \boldsymbol{e}_i$$

where we are again helped by $\dot{\boldsymbol{e}}_i = 0$. In the frame S' the expression is more complicated. From the expression $\dot{\boldsymbol{r}} = \dot{r}_i' \boldsymbol{e}_i' + \boldsymbol{\omega} \times \boldsymbol{r}$ with $\dot{\boldsymbol{e}}_i' = \boldsymbol{\omega} \times \boldsymbol{e}_i'$ and $\boldsymbol{r} = r_i' \boldsymbol{e}_i'$, time differentiation and a few steps of algebra lead to

$$\ddot{\boldsymbol{r}} = \ddot{r}_i' \boldsymbol{e}_i' + 2\dot{r}_i' \boldsymbol{\omega} \times \boldsymbol{e}_i' + r_i' \dot{\boldsymbol{\omega}} \times \boldsymbol{e}_i' + r_i' \boldsymbol{\omega} \times (\boldsymbol{\omega} \times \boldsymbol{e}_i')$$

- We adopt a analogous notation for acceleration as we did for velocities: the acceleration of the particle seen by an observer in S is

$$\left(\frac{\mathrm{d}^2 \boldsymbol{r}}{\mathrm{d}t^2} \right)_S = \ddot{r}_i \boldsymbol{e}_i$$

and the acceleration as seen by an observer rotating with the frame S' is

$$\left(\frac{\mathrm{d}^2 \boldsymbol{r}}{\mathrm{d}t^2} \right)_{S'} = \ddot{r}_i' \boldsymbol{e}_i'$$

If we equate the expressions for $\ddot{\boldsymbol{r}}$ in the inertial and rotating bases and adopt the new notation, leaving the quantities $\boldsymbol{r}$ and $\boldsymbol{\omega}$ in vector form (i.e. not expanding them in either the $\boldsymbol{e}_i$ or $\boldsymbol{e}_i'$ bases) we get the vector equation

$$\left(\frac{\mathrm{d}^2 \boldsymbol{r}}{\mathrm{d}t^2} \right)_S = \left(\frac{\mathrm{d}^2 \boldsymbol{r}}{\mathrm{d}t^2} \right)_{S'} + 2\boldsymbol{\omega} \times \left(\frac{\mathrm{d}\boldsymbol{r}}{\mathrm{d}t} \right)_{S'} + \boldsymbol{\omega} \times (\boldsymbol{\omega} \times \boldsymbol{r}) + \dot{\boldsymbol{\omega}} \times \boldsymbol{r}$$

This result is important: it lets us generalize Newton's second law to rotating frames.

1.3.2   Newton's Second Law in Rotating Frames

- In the inertial frame, using the new notation, Newton's second law $m\ddot{\boldsymbol{r}} = \boldsymbol{F}$ reads

$$m\left(\frac{\mathrm{d}^2\boldsymbol{r}}{\mathrm{d}t^2}\right)_S = \boldsymbol{F}$$

where $\boldsymbol{F}$ is the net force acting on the particle. Combining this with our last result, multiplying through by m and rearranging produces

$$m\left(\frac{\mathrm{d}^2\boldsymbol{r}}{\mathrm{d}t^2}\right)_S' = \boldsymbol{F} - 2m\boldsymbol{\omega} \times \left(\frac{\mathrm{d}\boldsymbol{r}}{\mathrm{d}t}\right)_{S'} - m\boldsymbol{\omega} \times (\boldsymbol{\omega} \times \boldsymbol{r}) - m\dot{\boldsymbol{\omega}} \times \boldsymbol{r}$$

This is Newton's second law in the rotating frame S'. To explain the particle's motion, an observer rotating with S' must introduce the existence of three further terms on the right side of the equation besides the traditional net force $\boldsymbol{F}$. The three extra terms are called *fictitious forces* and are a legitimate physical consequence of the frame S''s rotation.

From the perspective of an observer rotating with S' a free particle (for which $\boldsymbol{F} = 0$) does not travel with uniform motion $m\left(\frac{\mathrm{d}^2\boldsymbol{r}}{\mathrm{d}t^2}\right)_S' = 0$ because of the three fictitious forces.

1.3.3   Fictitious Forces

We will now analyze each term in the generalized second law

$$m\left(\frac{\mathrm{d}^2\boldsymbol{r}}{\mathrm{d}t^2}\right)_S' = \boldsymbol{F} - 2m\boldsymbol{\omega} \times \left(\frac{\mathrm{d}\boldsymbol{r}}{\mathrm{d}t}\right)_{S'} - m\dot{\boldsymbol{\omega}} \times \boldsymbol{r} - m\boldsymbol{\omega} \times (\boldsymbol{\omega} \times \boldsymbol{r})$$

1. The term $-2m\boldsymbol{\omega} \times \left(\frac{\mathrm{d}\boldsymbol{r}}{\mathrm{d}t}\right)_{S'}$ is called the *Coriolis force*.

   - It depends on the velocity $\left(\frac{\mathrm{d}\boldsymbol{r}}{\mathrm{d}t}\right)_{S'}$ of the particle as measured by an observer rotating with S'. Because the force is velocity dependent it is felt only by moving particles. Note also that the force is independent of position.

   - The Coriolis force has an interesting effect: it makes moving particles turn in circles. In fact, you probably (sort of) already knew this: the Coriolis force is mathematically identical to the familiar Lorentz force felt by a charged particle moving in a magnetic field. Just compare the equations:

   $$\boldsymbol{F}_{\mathrm{cor}} \sim m\boldsymbol{v} \times \boldsymbol{\omega} \qquad \text{and} \qquad \boldsymbol{F}_{\mathrm{mag}} \sim q\boldsymbol{v} \times \boldsymbol{B}$$

   We know the Lorentz force makes moving charged particles turn in circles and can expect the Coriolis force to do the same. More formally, the circles result because the force is always perpendicular to velocity by virtue of the cross product $\boldsymbol{v} \times \boldsymbol{\omega}$.

   - The magnitude of the Coriolis force is

   $$|\boldsymbol{F}_{\mathrm{cor}}| = 2m\omega v \sin\theta$$

   where θ is the angle between the particle's velocity $\boldsymbol{v}$ and the axis of rotation $\boldsymbol{\omega}$.

2. The term $-m\boldsymbol{\omega} \times (\boldsymbol{\omega} \times \boldsymbol{r})$ is called the *centrifugal force*.

   - With the vector identity $\boldsymbol{a} \times (\boldsymbol{b} \times \boldsymbol{c}) = (\boldsymbol{a} \cdot \boldsymbol{c})\boldsymbol{b} - (\boldsymbol{a} \cdot \boldsymbol{b})\boldsymbol{c}$ it can be written

     $$\begin{aligned} \boldsymbol{F}_{\text{cent}} &= -m\boldsymbol{\omega} \times (\boldsymbol{\omega} \times \boldsymbol{r}) \\ &= -m(\boldsymbol{\omega} \cdot \boldsymbol{r})\boldsymbol{\omega} + m\omega^2 \boldsymbol{r} \end{aligned}$$

     The centrifugal force on a particle points from the particle perpendicularly away from the axis of rotation.

   - The magnitude of the centrifugal force is

     $$|\boldsymbol{F}_{\text{cent}}| = m\omega^2 \operatorname{proj}_{\boldsymbol{\omega}}(\boldsymbol{r})$$

     where $\operatorname{proj}_{\boldsymbol{\omega}}(\boldsymbol{r})$ is the orthogonal projection of the position vector onto the axis of rotation $\boldsymbol{\omega}$.

   - There's an interesting interpretation of the centrifugal force. First, note that the force does not depend on the particle's velocity, only on its position $\boldsymbol{r}$. This suggest we can view the centrifugal force as a conservative force, with an associated potential, the form

     $$\boldsymbol{F}_{\text{cent}} = -\nabla V(\boldsymbol{r}) \qquad \text{where} \qquad V(\boldsymbol{r}) = -\frac{m}{2}|\boldsymbol{\omega} \times \boldsymbol{r}|^2$$

     In the rotating frame V has the interpretation of a potential energy depending on the particle's position associated with the frame's rotation. The potential is negative. We know particles tend to minimize their energy, so we expect particles in a centrifugal potential to fly out from the axis of rotation; increasing $|r|$ will minimize the potential energy. This agrees with our earlier conclusion that that centrifugal force acts outward from the axis of rotation.

3. The term $-m\dot{\boldsymbol{\omega}} \times \boldsymbol{r}$ is called the *Euler force*. It arises only when S' rotates with variable angular velocity $\boldsymbol{\omega}$ (i.e. $\dot{\boldsymbol{\omega}} \neq 0$). Because most everyday examples of non-inertial frames rotate with (nearly) constant angular velocity, the Euler force is usually negligible.

2   Lagrangian Mechanics

2.1   Constraints and Generalized Coordinates

Explain the role of constraints and generalized coordinates in Lagrangian mechanics.

2.1.1   Holonomic Constraints and Generalized Coordinates

- To describe a system of N particles in three dimensions we begin with $3N$ coordinates x^A where $A = 1, \ldots 3N$, since $3N$ is the default number of coordinates needed to describe N particles in 3 dimensions. The system is said to have $3N$ degrees of freedom.

- Formally, constraints are conditions that reduce a system's degrees of freedom. In Lagrangian mechanics constraints are often desirable since they make a system's motion simpler. A *holonomic constraint* is a relationship between the coordinates x^A that can be written in the form

$$f_\alpha(x_A, t) = 0 \qquad \alpha = 1, \ldots, 3N - n$$

  where t allows for the constraint to depend on time. A system as a whole is called a *holonomic system* if all of its constraints are holonomic.

- For the sake of completeness, a few words on something we won't have to deal with in this course. But it's good to know such things exist.

  Constraints that cannot be written in the form $f(x_A, t) = 0$ are called *non-holonomic* constraints. Two common examples are: constraints with inequalities and velocity-dependent constraints that cannot be integrated into the form $f(x^A, t) = 0$, e.g;

$$\begin{aligned}
\text{Constraint with an inequality:} \quad & f(x_A, t) \leq C, \quad (C \in \mathbb{R}) \\
\text{Velocity dependent constraint:} \quad & g(x^A, \dot{x}^A, t) = 0
\end{aligned}$$

  And a some vocabulary: If a constraint has an explicit time dependence, it is called *rheonomic*. If a constraint is not explicitly time-dependent, it is called *scleronomic*

- A system with $3N - n$ holonomic constraints can solved in terms of n *generalized coordinates* $q_1, \ldots, q_n$ to which we will assign the index $i = i, \ldots, n$ and label q_i. Once we solve a system's constraints in terms of the generalized coordinates q_i we can transform from the coordinates $x^A, A = 1, \ldots, 3N$ to the generalized coordinates q_i vis

$$x_A = x_A(q_1, \ldots, q_n)$$

  A system with $3N - n$ constraints and thus n generalized coordinates has n degrees of freedom, since we need n generalized coordinates to uniquely describe the system.

2.1.2   Lagrange Multipliers

- For each of the $3N - n$ constraints f_α we introduce $3N - n$ new variables λ_α, called *Lagrange multipliers*. In terms of the Lagrange multipliers we define a new Lagrangian

$$L' = L(x^A, \dot{x}^A, t) - \lambda_\alpha f_\alpha(x^A, t)$$

  We can treat the Lagrange multipliers λ_α as new coordinates.

13

- Using the new Lagrangian L' the Lagrange-Euler equations for λ_α are then

$$\frac{\partial L'}{\partial \lambda_\alpha} - \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L'}{\partial \dot\lambda_\alpha}\right) = 0 \implies f_\alpha(x^A, t) = 0$$

This simply recovers the constraints $f_\alpha(x^A, t) = 0$.

- The Lagrangian equations for x^A are more interesting. We have

$$\frac{\partial L'}{\partial x^A} - \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L'}{\partial \dot x^A}\right) = 0 \implies \frac{\partial L}{\partial x^A} + \lambda_\alpha \frac{\partial f_\alpha}{\partial x^A} - \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot x^A}\right) = 0$$

In other words, the Lagrangian formalism describes constrained systems with the equation

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot x^A}\right) - \frac{\partial L}{\partial x^A} = \lambda_\alpha \frac{\partial f_\alpha}{\partial x^A}$$

The left hand side is the Lagrange equation of motion for the unconstrained system, while the right hand side is the manifestation of the constraint forces.

2.1.3   Lagrange Equations in Generalized Coordinates

- The Lagrangian formalism makes it possible to solve for the motion of a constrained system in terms of the n generalized coordinates q_i instead of $3N$ coordinates x^A. This is really useful because n is usually small and $3N$ is usually large: it's easier to solve a few equations instead of many equations.

- Deriving the system's equations of motion directly in the generalized coordinates means it must be possible to write the Lagrangian in terms of the q_i in the form:

$$L(q_i, \dot q_i, t) = L\left(x^A(q_i, t), \dot x^A(q_i, \dot q_i, t)\right)$$

Here's a derivation showing this is possible. We'll work with the modified Lagrangian

$$L' = L(x^A, \dot x^A) - \lambda_\alpha f_\alpha(x^A, t)$$

and change from the x^A coordinates to the generalized coordinates and Lagrange multipliers:

$$x^A \to \begin{cases} q_i, & i = 1, \ldots, n \\ \lambda_\alpha, & \alpha = 1, \ldots, 3N \end{cases}$$

Using the Lagrangian $L' = L'(q_i, \dot q_i, \lambda_\alpha, t)$, the Lagrange equations for q_i are

$$\frac{\partial L'}{\partial q_i} - \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L'}{\partial \dot q_i}\right) = 0 \implies \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot q_i}\right) - \frac{\partial L}{\partial q_i} = \lambda_\alpha \frac{\partial f_\alpha}{\partial q_i}$$

But because the constraints f_α are by definition independent of the generalized coordinates q_i, we have $\lambda_\alpha \frac{\partial f_\alpha}{\partial q_i} = 0$ leaving

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot q_i}\right) - \frac{\partial L}{\partial q_i} = 0$$

showing we can derive the equations of motion directly from the Lagrangian $L(q_i, \dot q_i, t)$.

2.1.4   A Summary of Constraints and Coordinates in Lagrangian Mechanics

For a system of N particles in three dimensions:

- In Lagrangian mechanics we begin with $3N$ coordinates x^A; $3N$ is the default number of coordinates needed to describe N particles in 3 dimensions.

  Usually the system has some constraints. These simplify the motion and allow us to describe the dynamics with fewer coordinates, called generalized coordinates.

  For each additional constraint, we need one fewer coordinate to describe our system. If we have $3N - n$ constraints, we then need only $3N - (3N - n) = n$ generalized coordinates. Basically, the generalized coordinates are the simplified coordinates left over after applying all the constraints.

- Because it's fewer variable to work with, we describe a system in terms of the n generalized coordinates q_i with $i = 1, \ldots, n$. These coordinates form an n-dimensional space called the configuration space.

  The system's time evolution corresponds to a n-dimensional curve in configuration space. The time evolution is governed by the Lagrangian

  $$L = L(q_i, \dot{q}_i, t) = T(\dot{q}_i) - V(q_i)$$

  such that the coordinates q_i oby the Lagrange equations

  $$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{q}_i}\right) - \frac{\partial L}{\partial q_i} = 0$$

  where a system of n second order (generally) non-linear differential equations. The goal of Lagrangian mechanics is then solving the Lagrange equations for the dynamics of the generalized coordinates q_i.

2.2   Virtual Displacements and d'Alembert's Principle

Explain virtual displacements and d'Alembert's principle.

- Formally, à la Goldstein, a *virtual displacement* of a system is a change in the system's configuration as the result of an arbitrary infinitesimal change of the coordinates *at a given instant in time* consistent with the forces and constraints acting on the system at that instant. The displacement is called virtual to distinguish it from an actual displacement in a time *interval* $\mathrm{d}t$, during which the the forces and constraints might change.

  To distinguish them from displacements $\mathrm{d}\boldsymbol{r}$ than occur in a finite time interval $\mathrm{d}t$, we denote virtual displacements by $\delta\boldsymbol{r}$.

  Virtual work is the work of a force in the direction of a virtual displacement

  $$\delta W = \boldsymbol{F} \cdot \delta\boldsymbol{r}$$

- To get to d'Alembert's principle we first split the forces acting on our system into active forces and constraint forces.

  $$\boldsymbol{F}_i = \boldsymbol{F}_i^{(a)} + \boldsymbol{F}_i^{(c)}$$

15

Active forces are the forces that directly affect the dynamics of the generalized co-ordinates, while the constraint forces work passively to restrict the system's possible configurations.

Constraint forces that are perpendicular to the system's motion for all displacements, including virtual displacements, are called *ideal constraints*. These are the common constraints like normal force and tension. For a system with ideal constraints, in which the constraint forces are always perpendicular to motion, the virtual work done by constraint forces is zero.

$$\delta W^{(c)} = \boldsymbol{F}^{(c)} \cdot \delta \boldsymbol{r} = 0$$

- For a system in equilibrium, the net force on each particle vanishes, so $\boldsymbol{F}_i = 0$, and the virtual work done by the net force also vanishes. We then have

$$\boldsymbol{F}_i^{(\mathrm{eq})} \cdot \delta \boldsymbol{r}_i = 0$$

where the notation $\boldsymbol{F}_i^{(\mathrm{eq})}$ stresses the net force is such that the system is in equilibrium. With the decomposition of the net force into $\boldsymbol{F}_i^{(\mathrm{eq})} = \boldsymbol{F}_i^{(a)} + \boldsymbol{F}_i^{(c)}$, this becomes

$$\left(\boldsymbol{F}_i^{(a)} + \boldsymbol{F}_i^{(c)} \right) \cdot \delta \boldsymbol{r}_i = 0$$

For a system with ideal constraints the virtual work done by the constraint forces vanishes, leaving

$$\boldsymbol{F}_i^{(a)} \cdot \delta \boldsymbol{r}_i = 0$$

for any system in equilibrium. This result is called the *principle of virtual work*: for any system in equlibrium with ideal constraints, the total virtual work done by the active forces must vanish.

Note that this condition requires only the *total* virtual work is zero, so the *individual* $\boldsymbol{F}_i^{(a)}$ need not be zero.

- For a dynamic system, we use a little trick. We take Newton's equation of motion $\boldsymbol{F}_i = \dot{\boldsymbol{p}}_i$ and rearrange it to get

$$\boldsymbol{F}_i^{(\mathrm{eq})} \equiv \boldsymbol{F}_i - \dot{\boldsymbol{p}}_i = 0$$

which we interpret, à la Goldstein, as follows: the accelerating system would be in equilibrium if acted on by a net force $\boldsymbol{F}_i^{(\mathrm{eq})}$ equal to the accelerating force $\boldsymbol{F}_i$ minus a "reversed effective force" $\dot{\boldsymbol{p}}_i$. We can then plug $\boldsymbol{F}_i^{(\mathrm{eq})} = (\boldsymbol{F}_i - \dot{\boldsymbol{p}}_i)$ into the equilibrium equation $\boldsymbol{F}_i^{(\mathrm{eq})} \cdot \delta \boldsymbol{r}_i = 0$ to get

$$(\boldsymbol{F}_i - \dot{\boldsymbol{p}}_i) \cdot \delta \boldsymbol{r}_i = 0$$

With the decomposition $\boldsymbol{F}_i = \boldsymbol{F}_i^{(a)} + \boldsymbol{F}_i^{(c)}$ we get, written out in unnecessary detail:

$$0 = \left[\left(\boldsymbol{F}_i^{(a)} + \boldsymbol{F}_i^{(c)} \right) - \dot{\boldsymbol{p}}_i \right] \cdot \delta \boldsymbol{r}_i = \left(\boldsymbol{F}_i^{(a)} - \dot{\boldsymbol{p}}_i \right) \cdot \delta \boldsymbol{r}_i + \boldsymbol{F}_i^{(c)} \cdot \delta \boldsymbol{r}_i$$
$$= \left(\boldsymbol{F}_i^{(a)} - \dot{\boldsymbol{p}}_i \right) \cdot \delta \boldsymbol{r}_i + 0$$

The result

$$\left(\boldsymbol{F}_i^{(a)} - \dot{\boldsymbol{p}}_i \right) \cdot \delta \boldsymbol{r}_i = 0$$

is called *D'Alembert's principle*. The result is important because it applies to dynamics as well as statics, and only contains the active forces. I stress again that the quantity $\boldsymbol{F}_i^{(a)} - \dot{\boldsymbol{p}}_i$ is *not* in general zero for each particle individually, just that the total $\boldsymbol{F}^{(a)} - \dot{\boldsymbol{p}}$ for the particles together sums to zero.

2.3   Euler-Lagrange Equations à la d'Alembert

Derive the Euler-Lagrange equations from d'Alembert's principle.

- We start with D'Alembert's principle (written with the index $j = 1, \ldots, N$)

$$\left(\boldsymbol{F}_j^{(a)} - \dot{\boldsymbol{p}}_j \right) \cdot \delta \boldsymbol{r}_j = 0$$

D'Alembert's principle in its current form can't produce well-defined equations of motion because the coordinates $\boldsymbol{r}_j$ are in general dependent on each other. This is the same reason why the individual $\boldsymbol{F}_i^{(a)} - \dot{\boldsymbol{p}}_i$ are not zero.

To derive Lagrange's equations of motion, we need to transform D'Alembert's principle from $\boldsymbol{r}_j$ to the generalized coordinates q_i, which are independent.

- We start with the transformation between $\boldsymbol{r}_j$ and q_i:

$$\boldsymbol{r}_j = \boldsymbol{r}_j(q_1, \ldots, q_n, t) \qquad j = 1, \ldots, N$$

The total derivative of $\boldsymbol{r}_j$ is then

$$\mathrm{d}\boldsymbol{r}_j = \frac{\partial \boldsymbol{r}_j}{\partial q_i} \, \mathrm{d}q_i + \frac{\partial \boldsymbol{r}_j}{\partial t} \, \mathrm{d}t$$

For virtual displacements, which by definition occur for $\mathrm{d}t = 0$, we have

$$\delta \boldsymbol{r}_j = \frac{\partial \boldsymbol{r}_j}{\partial q_i} \delta q_i$$

With the expression for $\delta \boldsymbol{r}_j$, the total virtual work $\boldsymbol{F}_j \cdot \delta \boldsymbol{r}_j$ becomes

$$\boldsymbol{F}_j \cdot \delta \boldsymbol{r}_j = \boldsymbol{F}_j \cdot \left(\frac{\partial \boldsymbol{r}_j}{\partial q_i} \delta q_i \right) \equiv Q_i \delta q_i$$

where we defined a *generalized force* with components

$$Q_i = \boldsymbol{F}_j \cdot \frac{\partial \boldsymbol{r}_j}{\partial q_i}$$

For an system with ideal constraints, the constraint force do no work, so

$$Q_i = \boldsymbol{F}_j \cdot \frac{\partial \boldsymbol{r}_j}{\partial q_i} = \left(\boldsymbol{F}_j^{(a)} + \boldsymbol{F}_j^{(c)} \right) \cdot \frac{\partial \boldsymbol{r}_j}{\partial q_i} = \boldsymbol{F}_j^{(a)} \cdot \frac{\partial \boldsymbol{r}_j}{\partial q_i}$$

With the transformation $\boldsymbol{F}_j^{(a)} \cdot \delta \boldsymbol{r}_j = Q_i \delta q_i$ the first term in D'Alembert's principle is written purely in terms of independent generalized quantities.

- The transformation from $\boldsymbol{r}_j$ to q_i in the second term $\dot{\boldsymbol{p}}_j \cdot \delta \boldsymbol{r}_j$ requires a bit more effort. Writing $\dot{\boldsymbol{p}} = m\ddot{\boldsymbol{r}}$ and substituting the expression for $\delta \boldsymbol{r}_j$, we get

$$\dot{\boldsymbol{p}}_j \cdot \delta \boldsymbol{r}_j = m_j \ddot{\boldsymbol{r}}_j \cdot \left(\frac{\partial \boldsymbol{r}_j}{\partial q_i} \delta q_i \right)$$

By reverse engineering the product rule, we can write

$$m_j \ddot{\boldsymbol{r}}_j \frac{\partial \boldsymbol{r}_j}{\partial q_i} = \frac{\mathrm{d}}{\mathrm{d}t} \left[m_j \dot{\boldsymbol{r}}_j \frac{\partial \boldsymbol{r}_j}{\partial q_i} \right] - \dot{\boldsymbol{r}}_j \frac{\mathrm{d}}{\mathrm{d}t} \left[m_j \frac{\partial \boldsymbol{r}_j}{\partial q_i} \right]$$

To get an expression for $\dot{\boldsymbol{r}}_j$, differentiate $\boldsymbol{r}_j = \boldsymbol{r}_j(q_1, \ldots, q_N, t)$ with the chain rule to get

$$\dot{\boldsymbol{r}}_j = \frac{\mathrm{d}\boldsymbol{r}_j}{\mathrm{d}t} = \frac{\partial \boldsymbol{r}_j}{\partial q_i} \dot{q}_i + \frac{\partial \boldsymbol{r}_j}{\partial t}$$

17

- The next step is basically the foundation of the Lagrangian formalism (i.e. it's important): we define the $\dot{q}_i$ as a *new variable, independent of the q_i*. Even though they both involve the coordinates q, the Lagrangian formalism treats the q_i and $\dot{q}_i$ as independent variables.

  If we can view the $\dot{q}_i$ as new, independent coordinates, we can take the partial derivative of $\dot{\boldsymbol{r}}_j = \frac{\partial \boldsymbol{r}_j}{\partial q_i}\dot{q}_i + \frac{\partial \boldsymbol{r}_j}{\partial t}$ with respect to $\dot{q}_i$ with the product rule to get

  $$\frac{\partial \dot{\boldsymbol{r}}_j}{\partial \dot{q}_i} = \frac{\partial}{\partial \dot{q}_i}\left[\frac{\partial \boldsymbol{r}_j}{\partial q_i}\right]\dot{q}_i + \frac{\partial \boldsymbol{r}_j}{\partial q_i}\frac{\partial \dot{q}_i}{\partial \dot{q}_i} + \frac{\partial}{\partial \dot{q}_i}\left[\frac{\partial \boldsymbol{r}_j}{\partial t}\right]$$
  $$= 0 + \frac{\partial \boldsymbol{r}_j}{\partial q_i} + 0 \implies \frac{\partial \dot{\boldsymbol{r}}_j}{\partial \dot{q}_i} = \frac{\partial \boldsymbol{r}_j}{\partial q_i}$$

  The result $\frac{\partial \dot{\boldsymbol{r}}_j}{\partial \dot{q}_i} = \frac{\partial \boldsymbol{r}_j}{\partial q_i}$ looks like we just trivially "canceled the dots" but we could only do that because we defined the $\dot{q}_i$ as independent coordinates.

- Back to where we left off with

  $$m_j \ddot{\boldsymbol{r}}_j \frac{\partial \boldsymbol{r}_j}{\partial q_i} = \frac{\mathrm{d}}{\mathrm{d}t}\left[m_j \dot{\boldsymbol{r}}_j \frac{\partial \boldsymbol{r}_j}{\partial q_i}\right] - \dot{\boldsymbol{r}}_j \frac{\mathrm{d}}{\mathrm{d}t}\left[m_j \frac{\partial \boldsymbol{r}_j}{\partial q_i}\right] \equiv \mathrm{I} - \mathrm{II}$$

  where we'll tackle the expression term by term. Starting with I, we'll use the result $\frac{\partial \dot{\boldsymbol{r}}_j}{\partial \dot{q}_i} = \frac{\partial \boldsymbol{r}_j}{\partial q_i}$ and a reverse engineered chain rule to get

  $$\mathrm{I} \equiv \frac{\mathrm{d}}{\mathrm{d}t}\left[m_j \dot{\boldsymbol{r}}_j \frac{\partial \boldsymbol{r}_j}{\partial q_i}\right] = \frac{\mathrm{d}}{\mathrm{d}t}\left[m_j \dot{\boldsymbol{r}}_j \frac{\partial \dot{\boldsymbol{r}}_j}{\partial \dot{q}_i}\right] = \frac{\mathrm{d}}{\mathrm{d}t}\left[\frac{\partial}{\partial \dot{q}_i}\left(\frac{m_j \dot{\boldsymbol{r}}_j^2}{2}\right)\right] = \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial T}{\partial \dot{q}_i}\right)$$

- On to the second term $\mathrm{II} \equiv m_j \dot{\boldsymbol{r}}_j \frac{\mathrm{d}}{\mathrm{d}t}\left[\frac{\partial \boldsymbol{r}_j}{\partial q_i}\right]$. First, we'll calculate the derivative $\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial \boldsymbol{r}_j}{\partial q_i}$ using the chain rule, remembering $\boldsymbol{r}_j = \boldsymbol{r}_j(q_1, \ldots, q_n, t)$:

  $$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial \boldsymbol{r}_j}{\partial q_i} = \frac{\partial^2 \boldsymbol{r}_j}{\partial q_i \partial q_k}\dot{q}_k + \frac{\partial^2 \boldsymbol{r}_j}{\partial q_i \partial t} = \frac{\partial}{\partial q_i}\left[\frac{\partial \boldsymbol{r}_j}{\partial q_k}\dot{q}_k + \frac{\partial \boldsymbol{r}_j}{\partial t}\right] = \frac{\partial}{\partial q_i}\left[\frac{\mathrm{d}\boldsymbol{r}_j}{\mathrm{d}t}\right] = \frac{\partial \dot{\boldsymbol{r}}_j}{\partial q_i}$$

  where we can factor out the derivative $\frac{\partial}{\partial q_i}$ in the second equality because the q_i and $\dot{q}_i$ are different variables, so $\frac{\partial \dot{q}_k}{\partial q_i} = 0$ for all i, k.

  With the result $\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial \boldsymbol{r}_j}{\partial q_i} = \frac{\partial \dot{\boldsymbol{r}}_j}{\partial q_i}$ and a reverse engineered chain rule we get

  $$\mathrm{II} = m_j \dot{\boldsymbol{r}}_j \frac{\mathrm{d}}{\mathrm{d}t}\left[\frac{\partial \boldsymbol{r}_j}{\partial q_i}\right] = m_j \dot{\boldsymbol{r}}_j \frac{\partial \dot{\boldsymbol{r}}_j}{\partial q_i} = \frac{\partial}{\partial q_i}\left[\frac{m_j \dot{\boldsymbol{r}}_j^2}{2}\right] = \frac{\partial T}{\partial q_i}$$

- Putting the pieces together, D'Alembert's principle in terms of q_i reads

  $$\left[\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial T}{\partial \dot{q}_i}\right) - \frac{\partial T}{\partial q_i} - Q_j\right]\delta q_i = 0$$

  Here's a summary of how we put the pieces back together, starting with D'Alembert's principle $\left(\boldsymbol{F}_j^{(a)} - \dot{\boldsymbol{p}}_j\right) \cdot \delta \boldsymbol{r}_j = 0$.

  $$0 = \boldsymbol{F}_j^{(a)} \cdot \delta \boldsymbol{r}_j - \dot{\boldsymbol{p}}_j \cdot \delta \boldsymbol{r}_j = Q_i \delta q_i - m_j \ddot{\boldsymbol{r}}_j \cdot \frac{\partial \boldsymbol{r}_j}{\partial q_i}\delta q_i = [Q_i - (\mathrm{I} - \mathrm{II})]\delta q_i$$
  $$= \left[Q_i - \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial T}{\partial \dot{q}_i}\right) + \frac{\partial T}{\partial q_i}\right]\delta q_i \implies \left[\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial T}{\partial \dot{q}_i}\right) - \frac{\partial T}{\partial q_i} - Q_i\right]\delta q_i = 0$$

18

- Because the q_i *independent* of each other, the virtual displacements in separate coordinates δq_i and δq_j for $i \neq j$ are also independent, and the equality in D'Alembert's principle holds for *all* $i = 1, \ldots, n$ only if

$$\left[\frac{\mathrm{d}}{\mathrm{d}t} \left(\frac{\partial T}{\partial \dot{q}_i} \right) - \frac{\partial T}{\partial q_i} - Q_i \right] = 0 \qquad \text{or} \qquad \frac{\mathrm{d}}{\mathrm{d}t} \left(\frac{\partial T}{\partial \dot{q}_i} \right) - \frac{\partial T}{\partial q_i} = Q_i$$

which are a set of general equations of motion for Lagrangian mechanics.

- Usually we assume the forces acting on our system are conservative and thus can be derived from a scalar potential via

$$\boldsymbol{F} = -\nabla V(\boldsymbol{r}_1, \ldots, \boldsymbol{r}_N) \qquad \text{or, by components} \qquad \boldsymbol{F}_j = -\frac{\partial V}{\partial \boldsymbol{r}_j}$$

In this case, the generalized forces $Q_i = \boldsymbol{F}_j \cdot \frac{\partial \boldsymbol{r}_j}{\partial q_i}$ can be conveniently written

$$Q_i = \left(-\frac{\partial V}{\partial \boldsymbol{r}_j} \right) \cdot \frac{\partial \boldsymbol{r}_j}{\partial q_i} = -\frac{\partial V}{\partial q_i}$$

Plugging Q_i into the equations of motion then gives

$$\frac{\mathrm{d}}{\mathrm{d}t} \left(\frac{\partial T}{\partial \dot{q}_i} \right) - \frac{\partial T}{\partial q_i} - \left(-\frac{\partial V}{\partial q_i} \right) = \frac{\mathrm{d}}{\mathrm{d}t} \left(\frac{\partial T}{\partial \dot{q}_i} \right) - \frac{\partial (T - V)}{\partial q_i} = 0$$

- One more assumption: we restrict our system to velocity independent potentials (which still covers most of the common interactions in nature, e.g. gravitational, electrostatic, etc...). A velocity-independent potential means $V \neq V(\dot{q}_i)$. The immediate implication $\frac{\partial V}{\partial \dot{q}_i} = 0$ lets us rewrite

$$\frac{\partial T}{\partial \dot{q}_j} = \frac{\partial (T - V)}{\partial \dot{q}_j}$$

The equations of motion then read

$$\frac{\mathrm{d}}{\mathrm{d}t} \left(\frac{\partial T}{\partial \dot{q}_i} \right) - \frac{\partial (T - V)}{\partial q_i} = \frac{\mathrm{d}}{\mathrm{d}t} \left[\frac{\partial (T - V)}{\partial \dot{q}_i} \right] - \frac{\partial (T - V)}{\partial q_i} = 0$$

Or, in terms of the Lagrangian $L \equiv T - V$, the so-called *Lagrange equations*

$$\frac{\mathrm{d}}{\mathrm{d}t} \left(\frac{\partial L}{\partial \dot{q}_i} \right) - \frac{\partial L}{\partial q_i} = 0$$

2.4   Euler-Lagrange Equations and the Least Action Principle

What is the principle of least action? Use the least action principle to derive the Euler-Lagrange equations.

- We start with a system of $3N$ particles with the coordinates x^A, $A = 1, \ldots, 3N$ under the influence of a conservative force. The system's Lagrangian is

$$L(x^A, \dot{x}^A) = T(\dot{x}^A) - V(x^A)$$

where $T = \frac{1}{2} m_A (\dot{x}^A)^2$ is the system's total kinetic energy and $V(x^A)$ is the system's total potential energy.

- The system's time evolution corresponds to a $3N$ dimensional path $x(t)$ in the system's configuration space. An *action* is a property of a path $x(t)$ in configuration space, and is defined as

$$S[x(t)] = \int_{t_i}^{t_f} L(x(t), \dot{x}(t)) \, \mathrm{d}t$$

where $L(x, \dot{x}) = T(\dot{x}) - V(x)$ is the system's Lagrangian.

- For the principle of least action, we consider all smooth paths in the configuration space between the fixed endpoints x_i and x_f where $x(t_i) = x_i$ and $x(t_f) = x_f$.

  The *principle of least action states*:

  > The path taken by a system between two points in its configuration space is an extremum of the action S.

  Note that, formally, the extremum need not be a minimum: it could be a maximum, minimum, or saddle point, although minima tend to correspond to physical scenarios.

- Consider a smooth paths through configuration space between the fixed endpoints x_i^A and x_f^A where $x^A(t_i) = x_i^A$ and $x^A(t_f) = x_f^A$. We'll write the path $x^A(t)$. Then consider varying the path just slightly. Instead of $x^A(t)$ we have

$$x^A \to x^A(t) + \delta x^A(t)$$

where δx^A is small compared to x^A.

- Although the interior of the path can vary as $x^A \to x^A(t) + \delta x^A(t)$, we'll insist on keeping our endpoints fixed, so $x^A(t_i) = x_i^A$ and $x^A(t_f) = x_f^A$. In other words, the variation vanishes at the endpoints x_i^A and x_f^A and

$$\delta x^A(t_i) = \delta x^A(t_f) = 0 \implies x^A(t_i) = x_i^A \quad \text{and} \quad x^A(t_f) = x_f^A$$

If the talk of variation of paths in configuration space seems abstract, think in everyday terms of walking between the same two points x_i and x_f, but taking a slightly different path between them.

- The change in action δS corresponding to our slight change in path $\delta x^A(t)$ is

$$\delta S = \delta \left[\int_{t_i}^{t_f} L(x^A, \dot{x}^A) \, \mathrm{d}t \right] = \int_{t_i}^{t_f} \delta L(x^A, \dot{x}^A) \, \mathrm{d}t$$

$$= \int_{t_i}^{t_f} \left(\frac{\partial L}{\partial x^A} \delta x^A + \frac{\partial L}{\partial \dot{x}^A} \delta \dot{x}^A \right) \mathrm{d}t$$

$$= \int_{t_i}^{t_f} \frac{\partial L}{\partial x^A} \delta x^A \, \mathrm{d}t + \int_{t_i}^{t_f} \frac{\partial L}{\partial \dot{x}^A} \delta \dot{x}^A \, \mathrm{d}t$$

- We integrate the second integral by parts with

$$\mathrm{d}v = \delta \dot{x}^A \, \mathrm{d}t \implies v = \delta x^A \qquad \text{and} \qquad u = \frac{\partial L}{\partial \dot{x}^A} \implies \mathrm{d}u = \frac{\mathrm{d}}{\mathrm{d}t} \frac{\partial L}{\partial \dot{x}^A} \, \mathrm{d}t$$

Applying $\int u \, \mathrm{d}v = uv - \int v \, \mathrm{d}u$ and joining the two integrals into one integrand, the variation in action reads

$$\delta S = \int_{t_i}^{t_f} \left[\frac{\partial L}{\partial x^A} - \frac{\mathrm{d}}{\mathrm{d}t} \frac{\partial L}{\partial \dot{x}^A} \right] \delta x^A \, \mathrm{d}t + \left[\frac{\partial L}{\partial \dot{x}^A} \delta x \right]_{t_i}^{t_f}$$

Since we have fixed the endpoints, $\delta x^A(t_i) = \delta x^A(t_f) = 0$ and the last term (the *uv* term from integration by parts) vanishes. We're left with

$$\delta S = \int_{t_i}^{t_f} \left[\frac{\partial L}{\partial x^A} - \frac{\mathrm{d}}{\mathrm{d}t} \frac{\partial L}{\partial \dot{x}^A} \right] \delta x^A \,\mathrm{d}t$$

- Back to the principle of least action: the requirement that action S is an extremum holds only if $\delta S = 0$ for all small variations in the path $\delta x^A(t)$. Since $\delta x^A \neq 0$, the requirement $\delta S = 0$ for all paths holds if the other term in the integrand is zero, namely

$$\frac{\partial L}{\partial x^A} - \frac{\mathrm{d}}{\mathrm{d}t} \frac{\partial L}{\partial \dot{x}^A} = 0 \qquad \text{for all } A = 1, \ldots, 3N$$

These equations are called *Lagrange's equations*, and we'll see they replace Newton's equations in Lagrangian mechanics.

- To show that Lagrange's equations are equivalent to Newton's, we start by differentiating the Lagrangian $L(x^A, \dot{x}^A) = T(\dot{x}^A) - V(x^A)$ to get

$$\frac{\partial L}{\partial x^A} = -\frac{\partial V}{\partial x^A} \qquad \text{and} \qquad \frac{\partial L}{\partial \dot{x}^A} = \frac{\partial T}{\partial \dot{x}^A} \equiv \frac{\partial}{\partial \dot{x}^A} \left[\frac{1}{2} m \left(\dot{x}^A \right)^2 \right] = m\dot{x}^A = p_A$$

Putting the pieces together, we see the Lagrange and Newton equations are equivalent

$$\frac{\mathrm{d}}{\mathrm{d}t} \frac{\partial L}{\partial \dot{x}^A} = \frac{\partial L}{\partial x^A} \iff \dot{p}_A = -\frac{\partial V}{\partial x^A} \iff \dot{\boldsymbol{p}}_i = \boldsymbol{F}_i$$

2.5   Conserved Quantities, Symmetries and Noether's Theorem

What is the definition of a conserved quantity in Lagrangian mechanics? What is a continuous symmetry of the Lagrangian function? Give examples of each. State and prove (the simple version of) Noether's theorem, and provide examples of its application.

2.5.1   Constants of Motion and Some Examples

- First, a definition: for a system with the generalized coordinates q_i, a function $F(q_i, \dot{q}_i, t)$ (i.e. a function of the generalized coordinates, their time derivatives, and (possibly) time) is called a *constant of motion* or a *conserved quantity* if

$$\frac{\mathrm{d}F}{\mathrm{d}t} = \frac{\partial F}{\partial q_i} \dot{q}_i + \frac{\partial F}{\partial \dot{q}_i} \ddot{q}_i + \frac{\partial F}{\partial t} = 0$$

wherever $q_i(t)$ satisfy the Lagrange equations. In other words, its total time derivative must be zero for all allowed configurations of $q_i(t)$. Any conserved quantity F is constant along a path followed by the system through its configuration space.

- **First Example:** If L doesn't depend explicitly on time (i.e. $\frac{\partial L}{\partial t} = 0$), the quantity

$$H = \dot{q}_i \frac{\partial L}{\partial \dot{q}_i} - L$$

is conserved. To show the Hamiltonian is conserved when $L = L(q_i, \dot{q}_i)$ and $\frac{\partial L}{\partial t} = 0$ we calculate

$$\frac{\mathrm{d}H}{\mathrm{d}t} = \frac{\mathrm{d}}{\mathrm{d}t} \left[\dot{q}_i \frac{\partial L}{\partial \dot{q}_i} - L(q_i, \dot{q}_i) \right] = \ddot{q}_i \frac{\partial L}{\partial \dot{q}_i} + \dot{q}_i \frac{\mathrm{d}}{\mathrm{d}t} \left(\frac{\partial L}{\partial \dot{q}_i} \right) - \frac{\partial L}{\partial q_i} \dot{q}_i - \frac{\partial L}{\partial \dot{q}_i} \ddot{q}_i$$

Because $\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{q}_i}\right) = \frac{\partial L}{\partial q_i}$ wherever q_i satisfy the Lagrange equations we have

$$\frac{\mathrm{d}H}{\mathrm{d}t} = \ddot{q}_i \frac{\partial L}{\partial \dot{q}_i} - \frac{\partial L}{\partial \dot{q}_i}\ddot{q}_i + \dot{q}_i\left(\frac{\partial L}{\partial q_i}\right) - \frac{\partial L}{\partial q_i}\dot{q}_i = 0$$

H is called the system's *Hamiltonian* and is associated with the system's total energy. As long as $\frac{\partial V}{\partial \dot{q}_i} = 0$, $H = T + V$ is the system's total energy, which follows from

$$\dot{q}_i \frac{\partial}{\partial \dot{q}_i}\left[\frac{1}{2}m_i\dot{q}_i^2 + V(q_i)\right] = m_i\dot{q}_i^2 = 2T \implies \dot{q}_i \frac{\partial L}{\partial \dot{q}_i} - L = T + V$$

Second Example: If $\frac{\partial L}{\partial q_i} = 0$ for some q_i, then the quantity

$$p_i \equiv \frac{\partial L}{\partial \dot{q}_i}$$

is a constant of motion. The proof is a straightforward application of the Lagrange equations $\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{q}_i} = \frac{\partial L}{\partial q_i}$ and the condition $\frac{\partial L}{\partial q_i} = 0$:

$$\frac{\mathrm{d}p_i}{\mathrm{d}t} \equiv \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{q}_i}\right) = \frac{\partial L}{\partial q_i} = 0$$

The quantity p_i is called the *generalized momentum* corresponding to the generalized coordinate q_i. In fact, we can re-write the Lagrange equations in the form

$$\dot{p}_i = \frac{\partial L}{\partial q_i}$$

2.5.2 Symmetries of the Lagrangian

- Consider a one-parameter family of maps of the form

$$q_i(t) \to Q_i(s,t) \qquad \text{such that} \qquad Q_i(0,t) = q_i(t)$$

parameterized by $s \in \mathbb{R}$. Such a map is called a *continuous symmetry of the Lagrangian* $L = L(q_i, \dot{q}_i, t)$ if

$$\frac{\partial}{\partial s}L\left(Q_i(s,t), \dot{Q}_i(s,t), t\right) = 0 \qquad \text{for all } s \in \mathbb{R}$$

That probably sounds abstract. Basically, we want to show that the Lagrangians for the coordinates q_i and Q_i are the same, i.e. $L(q_i, \dot{q}_i, t) = L(Q_i, \dot{Q}_i, t)$.

- As an example, consider the symmetry of translation: the Lagrangian is the same under a translation of spatial coordinates. We write the map as

$$\boldsymbol{r}_i \to \boldsymbol{r}_i + s\hat{\mathbf{n}}$$

for some unit vector $\hat{\mathbf{n}}$ and scalar parameter s. This map is just a translation of the origin by the constant vector $s\hat{\mathbf{n}}$. In the language of q_i and Q_i, we have

$$q_i \leftrightarrow \boldsymbol{r}_i \qquad \text{and} \qquad Q_i \leftrightarrow \boldsymbol{r}_i + s\hat{\mathbf{n}}$$
$$\dot{q}_i \leftrightarrow \dot{\boldsymbol{r}}_i \qquad \text{and} \qquad \dot{Q}_i \leftrightarrow \dot{\boldsymbol{r}}_i$$

If the Lagrangian $L(q_i, \dot{q}_i, t)$ is defined as

$$L(q_i, \dot{q}_i, t) = \frac{1}{2} \sum_i m_i \dot{\boldsymbol{r}}_i^2 - V(|\boldsymbol{r}_i - \boldsymbol{r}_j|)$$

then the Lagrangian $L\big(Q_i(s,t), \dot{Q}_i(s,t), t\big) \leftrightarrow L(\boldsymbol{r}_i + s\hat{\mathbf{n}}, \dot{\boldsymbol{r}}_i, t)$ would be

$$L\big(Q_i(s,t), \dot{Q}_i(s,t), t\big) = \frac{1}{2} \sum_i m_i \dot{\boldsymbol{r}}_i^2 - V\big(|(\boldsymbol{r}_i + s\hat{\mathbf{n}}) - (\boldsymbol{r}_j + s\hat{\mathbf{n}})|\big)$$

$$= \frac{1}{2} \sum_i m_i \dot{\boldsymbol{r}}_i^2 - V(|\boldsymbol{r}_i - \boldsymbol{r}_j|)$$

It follows immediately that

$$\frac{\partial}{\partial s} L\big(Q_i(s,t), \dot{Q}_i(s,t), t\big) = 0 \qquad \text{and} \qquad L\big(Q_i(s,t), \dot{Q}_i(s,t), t\big) = L(q_i, \dot{q}_i, t)$$

showing by definition that space translation $\boldsymbol{r}_i \to \boldsymbol{r}_i + s\hat{\mathbf{n}}$ is a continuous symmetry of the Lagrangian. The second equation gives a more intuitive reason: the translated and untranslated Lagrangians are the same!

2.5.3 Statement of Noether's Theorem

Here it is:

For every continuous symmetry of the Lagrangian, there exists a corresponding constant of motion $\dfrac{\partial L}{\partial \dot{q}_i} \dfrac{\partial Q_i}{\partial s}\bigg|_{s=0}$

Note that this is not the most general version of Noether's theorem, which has a more powerful formulation in terms of infinitesimal variations of the coordinates q_i and time. Following is a proof of our simple version:

- Differentiating $L\left(Q_i(s,t), \dot{Q}_i(s,t), t\right)$ with the chain rule gives

$$\frac{\partial L}{\partial s} = \frac{\partial L}{\partial Q_i} \frac{\partial Q_i}{\partial s} + \frac{\partial L}{\partial \dot{Q}_i} \frac{\partial \dot{Q}_i}{\partial s}$$

By definition of a continuous symmetry $\frac{\partial L}{\partial s} = 0$ for all s. Choosing $s = 0$ and applying $Q_i\big|_{s=0} = q_i$ and $\dot{Q}_i\big|_{s=0} = \dot{q}_i$ to $\frac{\partial L}{\partial s}$ we have:

$$0 \equiv \frac{\partial L}{\partial s}\bigg|_{s=0} = \frac{\partial L}{\partial q_i} \frac{\partial Q_i}{\partial s}\bigg|_{s=0} + \frac{\partial L}{\partial \dot{q}_i} \frac{\partial \dot{Q}_i}{\partial s}\bigg|_{s=0}$$

- Next, we'll need the intermediate result $\frac{\partial \dot{Q}_i}{\partial s}\big|_{s=0} = \frac{\mathrm{d}}{\mathrm{d}t} \frac{\partial Q_i}{\partial s}\big|_{s=0}$. Recall $Q_i = Q_i(s,t)$. Working backwards and differentiating with the chain rule gives

$$\frac{\mathrm{d}}{\mathrm{d}t} \frac{\partial Q_i}{\partial s} = \frac{\partial^2 Q_i}{\partial s \partial t} \dot{s} + \frac{\partial^2 Q_i}{\partial s \partial t} = \frac{\partial}{\partial s}\left(\frac{\partial Q_i}{\partial s} \dot{s} + \frac{\partial Q_i}{\partial t}\right) = \frac{\partial}{\partial s}\left[\frac{\mathrm{d}Q_i}{\mathrm{d}t}\right] = \frac{\partial \dot{Q}_i}{\partial s}$$

- Back to the main proof, where we left off with

$$0 \equiv \frac{\partial L}{\partial s}\bigg|_{s=0} = \frac{\partial L}{\partial q_i}\frac{\partial Q_i}{\partial s}\bigg|_{s=0} + \frac{\partial L}{\partial \dot{q}_i}\frac{\partial \dot{Q}_i}{\partial s}\bigg|_{s=0}$$

Substituting $\frac{\partial L}{\partial q_i} = \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{q}_i}$ (the Lagrange equations), using $\frac{\partial \dot{Q}_i}{\partial s}\big|_{s=0} = \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial Q_i}{\partial s}\big|_{s=0}$ and recognizing a reverse product rule gives

$$0 \equiv \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{q}_i}\right)\frac{\partial Q_i}{\partial s}\bigg|_{s=0} + \frac{\partial L}{\partial \dot{q}_i}\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial Q_i}{\partial s}\right)\bigg|_{s=0} = \frac{\mathrm{d}}{\mathrm{d}t}\left[\frac{\partial L}{\partial \dot{q}_i}\frac{\partial Q_i}{\partial s}\right]_{s=0}$$

The result $\frac{\mathrm{d}}{\mathrm{d}t}\left[\frac{\partial L}{\partial \dot{q}_i}\frac{\partial Q_i}{\partial s}\right]_{s=0} = 0$ proves that $\frac{\partial L}{\partial \dot{q}_i}\frac{\partial Q_i}{\partial s}$ evaluated at $s = 0$ is indeed a constant of motion.

2.5.4 Examples with Noether's Theorem

Linear Momentum and the Homogeneity of Space

We'll briefly return to vector notation: it's easier for these examples.

- Consider an closed system of N particles with masses $m_1, \ldots, m_N$ and position vectors $\boldsymbol{r}_1, \ldots, \boldsymbol{r}_N$. The system's Lagrangian is

$$L = \frac{1}{2}\sum_i m_i \dot{\boldsymbol{r}}_i^2 - V(|\boldsymbol{r}_i - \boldsymbol{r}_j|)$$

- Recall that the symmetry of space translation is written

$$\boldsymbol{r}_i \to \boldsymbol{r}_i + s\hat{\mathbf{n}}$$

and that space translation $\boldsymbol{r}_i \to \boldsymbol{r}_i + s\hat{\mathbf{n}}$ is a valid continuous symmetry of the Lagrangian L since

$$\frac{\partial}{\partial s}L(\boldsymbol{r}_i + s\hat{\mathbf{n}}, \dot{\boldsymbol{r}}_i, t) = 0 \qquad \text{and} \qquad L(\boldsymbol{r}_i, \dot{\boldsymbol{r}}_i, t) = L(\boldsymbol{r}_i + s\hat{\mathbf{n}}, \dot{\boldsymbol{r}}_i, t)$$

This is a formal way of saying that space is homogeneous and a translation of the system by $s\hat{\mathbf{n}}$ does nothing to the equations of motion.

- Applying Noether's theorem to the symmetry $\boldsymbol{r}_i \to \boldsymbol{r}_i + s\hat{\mathbf{n}}$ tells us that

$$\frac{\partial L}{\partial \dot{q}_i}\frac{\partial Q_i}{\partial s}\bigg|_{s=0} = \sum_i^N \frac{\partial L}{\partial \dot{\boldsymbol{r}}_i}\cdot\hat{\mathbf{n}} = \sum_i^N \boldsymbol{p}_i \cdot \hat{\mathbf{n}}$$

is a constant of motion. The quantity $\sum_i \boldsymbol{p}_i \cdot \hat{\mathbf{n}}$ is the system's total linear momentum in the direction of the vector $\hat{\mathbf{n}}$. But since the choice of $\hat{\mathbf{n}}$ is arbitrary, $\sum_i \boldsymbol{p}_i \cdot \hat{\mathbf{n}}$ for all $\hat{\mathbf{n}}$, meaning the system's total linear momentum $\sum_i \boldsymbol{p}_i$ is conserved.

Angular Momentum and the Isotropy of Space

Isotropy of space is a fancy way of saying that space has the same properties in all directions. This manifests itself as invariance of space under rotations.

- Sticking with the closed system of N particles, if we rotate our system about some axis $\hat{\mathbf{n}}$, isotropy of space implies that all particles $\boldsymbol{r}_i \to \boldsymbol{r}_i'$ are rotated by the same amount. We can write the rotational mapping as an infinitesimal rotation

$$\boldsymbol{r}_i \to \boldsymbol{r}_i + \delta\theta\hat{\mathbf{n}} \times \boldsymbol{r}_i$$

  parameterized by the infinitesimal angle $\delta\theta$.

  Aside: We're working with infinitesimal rotations, which can be written in the form $\boldsymbol{r} \to \boldsymbol{r} + \delta\boldsymbol{r} = \boldsymbol{r} + \delta\theta\hat{\mathbf{n}} \times \boldsymbol{r}$ (instead of using a 3×3 rotation matrix). This makes the math easier. The justification is that we can reconstruct a finite rotation as a series of many infinitesimal rotations.

- If the Lagrangian is invariant under the rotational mapping, the conserved quantity is

$$\frac{\partial L}{\partial \dot{q}_i} \frac{\partial}{\partial \delta\theta} [\boldsymbol{r}_i + \delta\theta\hat{\mathbf{n}} \times \boldsymbol{r}_i] \Big|_{\delta\theta=0} = \sum_i^N \frac{\partial L}{\partial \dot{\boldsymbol{r}}_i} \cdot (\hat{\mathbf{n}} \times \boldsymbol{r}_i) = \sum_i^N \boldsymbol{p}_i \cdot (\hat{\mathbf{n}} \times \boldsymbol{r}_i)$$

$$= \sum_i^N \hat{\mathbf{n}} \cdot (\boldsymbol{r}_i \times \boldsymbol{p}_i) = \hat{\mathbf{n}} \cdot \sum_i^N \boldsymbol{L}_i = \hat{\mathbf{n}} \cdot \boldsymbol{L}$$

  which we recognize as the total angular momentum in the direction of $\hat{\mathbf{n}}$. The conclusion is the same as with linear momentum $\boldsymbol{p}$: since $\hat{\mathbf{n}} \cdot \boldsymbol{L}$ is conserved and $\hat{\mathbf{n}}$ is arbitrary, the system's total angular momentum $\boldsymbol{L}$ is conserved.

2.6   Lagrangian Formalism for a 1D Continuous Body

Explain the basics of the Lagrangian formalism for a continuous body. Include a derivation of the Lagrange equations for a one-dimensional continuous body.

2.6.1   Deriving the Lagrangian

- We model a one-dimensional continuous body as a series of $N + 1$ mass points separated by the distance a. The coordinate q_i of the ith point, where $i = 0, \ldots, N$, is

$$q_i(t) = ia + u_i(t) \equiv q_i^{(0)} + u_i(t)$$

  where $u_i(t)$ represents small deviations from the equilibrium position $q_i^{(0)}$. We assume the particles are connected as harmonic oscillators with potential energy $V_i = \frac{1}{2}k\Delta(q_i)^2$, where $\Delta q_i = u_{i+1} - u_i$ is the displacement from the equilibrium position and k is the spring constant. The system's Lagrangian is

$$L = \frac{1}{2}\sum_{i=0}^N m\dot{u}_i^2 - \frac{k}{2}\sum_{i=0}^N (u_{i+1} - u_i)^2 - \frac{k_b}{2}(u_0^2 + u_N^2)$$

  where the last term allows for boundary conditions with a different spring constant k_b. The Lagrange equations $\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{u}_i}\right) = \frac{\partial L}{\partial u_i}$ for the ith particle are

$$m\ddot{u}_i = k\left[(u_{i+1} - u_i) - (u_i - u_{i-1})\right] = k(u_{i+1} - 2u_i + u_{i-1})$$

  for $i = 1, \ldots, N - 1$.

- In the limit of a continuous substance with $N \to \infty$ and $a \to 0$ with $N \cdot a$ constant, we can approximate the right hand side as a second derivative using the finite difference approximation

$$\frac{\partial^2 f}{\partial x^2} \approx \frac{f(x + 2\Delta x) - 2f(x) + f(x + \Delta x)}{(\Delta x)^2}$$

The analogous expression is $\frac{u_{i+1} - 2u_i + u_{i+1}}{a^2} \approx \frac{\partial^2 u}{\partial x^2}$ and the Lagrange equations become

$$m\ddot{u} = ka^2 \frac{\partial^2 u}{\partial x^2} \implies \ddot{u} = \left(\frac{ka^2}{m} \right) \frac{\partial^2 u}{\partial x^2} = c^2 u_{xx}$$

where we've replaced $\frac{ka^2}{m}$ with the wave speed c in the substance and introduced the shorthand notation $\frac{\partial^2 u}{\partial x^2} \equiv u_{xx}$.

- Next, we replace energy quantities with energy densities in terms of the body's cross-sectional area S_0. The kinetic energy density is

$$\mathcal{T} = \frac{T}{S_0 a} = \frac{m}{2 S_0 a} u_t^2 = \frac{1}{2} \rho u_t^2$$

where ρ is the body's mass density. The potential energy density is

$$\mathcal{V} = \frac{V}{S_0 a} = \frac{1}{2} \frac{k(\Delta u_i)^2}{S_0 a} = \frac{1}{2} \frac{E_y (\Delta u_i)^2}{a^2} \approx \frac{1}{2} E_y u_x^2$$

where E_y is the body's Young's modulus. We approximated $\lim_{a \to 0} \frac{\Delta u_i}{a} \approx u_x$. The Lagrangian density is

$$\mathcal{L} = \mathcal{T} - \mathcal{V} = \frac{1}{2} \left(\rho u_t^2 - E_y u_x^2 \right)$$

In the limit $N \to \infty$ and $a \to 0$, the Lagrangian function in one dimension is

$$L = S_0 \int_0^l \mathcal{L}(u, u_x, t, u_t) \, \mathrm{d}x$$

where l is the length of the body.

2.6.2   Lagrange Equations for a 1D Continuous Substance

The Lagrange equations for a one dimensional continuous substance are

$$\frac{\mathrm{d}}{\mathrm{d}t} \left(\frac{\partial \mathcal{L}}{\partial t} \right) + \frac{\mathrm{d}}{\mathrm{d}x} \left(\frac{\partial \mathcal{L}}{\partial x} \right) - \frac{\partial L}{\partial u} = 0$$

We'll derive them using the principle of least action.

- Recall that the Lagrangian is

$$L = S_0$$

Action S for a physical process between two times t_1 and t_2 is defined as

$$S = \int_{t_1}^{t_2} L \, \mathrm{d}t = S_0 \int_{t_1}^{t_2} \int_0^l \mathrm{d}x \, \mathrm{d}t \mathcal{L}(u, u_x, t, u_t)$$

The principle of least action requires S is an extremum for all through the system's configuration space with fixed endpoints at t_1 and t_2. Because the coordinate $u = u(x, t)$ is a function of both x and t, we impose separate boundary conditions for both x and t, which imply the variation in δu vanishes at both the boundary points and boundary times.

$$u(x, t_1) = u(x, t_f) = 0 \qquad \text{and} \qquad u(0, t) = u(l, t) = 0$$
$$\delta u\big|_{t=t_i} = \delta u\big|_{t=t_f} = 0 \qquad \text{and} \qquad \delta u\big|_{x=0} = \delta u\big|_{x=l} = 0$$

- Applying the principle of least action and required $\delta S = 0$ leads to

$$0 \equiv \delta S = S_0 \int_{t_1}^{t_2} \int_0^l \mathrm{d}x\,\mathrm{d}t\,\delta\mathcal{L}(u, u_x, u_t, t)$$
$$= S_0 \int_{t_1}^{t_2} \int_0^l \mathrm{d}x\,\mathrm{d}t \left[\frac{\partial\mathcal{L}}{\partial u}\delta u + \frac{\partial\mathcal{L}}{\partial u_x}\delta u_x + \frac{\partial\mathcal{L}}{\partial u_t}\delta u_t\right]$$

We use integration by parts to simplify the second and third terms, integrating $\frac{\partial\mathcal{L}}{\partial u_x}\delta u_x$ over x and $\frac{\partial\mathcal{L}}{\partial u_t}\delta u_t$ over t.

$$u = \frac{\partial L}{\partial u_x} \implies \mathrm{d}u = \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial L}{\partial u_x}\right) \qquad \text{and} \qquad \mathrm{d}v = \delta u_x \implies v = \delta u$$
$$u = \frac{\partial L}{\partial u_t} \implies \mathrm{d}u = \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial u_t}\right) \qquad \text{and} \qquad \mathrm{d}v = \delta u_t \implies v = \delta u$$

The results are

$$\int_0^l \frac{\partial\mathcal{L}}{\partial u_x}\delta u_x = \left[\frac{\partial L}{\partial x}\delta u\right]_0^l - \int_0^l \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial L}{\partial u_x}\right)\delta u = -\int_0^l \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial L}{\partial u_x}\right)\delta u$$
$$\int_{t_1}^{t_2} \frac{\partial\mathcal{L}}{\partial u_t}\delta u_t = \left[\frac{\partial L}{\partial t}\delta u\right]_{t_1}^{t_2} - \int_{t_1}^{t_2} \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial u_t}\right)\delta u = -\int_{t_1}^{t_2} \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial u_t}\right)\delta u$$

where the uv terms in square brackets vanish because of the boundary conditions.

- With the results from integration by parts, we can rewrite the variation in action as

$$\delta S = S_0 \int_{t_1}^{t_2} \int_0^l \mathrm{d}x\,\mathrm{d}t \left[\frac{\partial\mathcal{L}}{\partial u} - \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial L}{\partial u_x}\right) - \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial u_t}\right)\right]\delta u$$

$\delta S = 0$ is satisfied for all paths only if the terms in the square brackets are zero, giving the Lagrange equations for a one-dimensional continuous substance:

$$\frac{\partial\mathcal{L}}{\partial u} - \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial L}{\partial u_x}\right) - \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial u_t}\right) = 0$$

2.6.3   Generalizations of Common Quantities For a Continuous Substance

Here is a very brief run-through of how the common quantities of Lagrangian mechanics are generalized to apply to a continuous substance.

- The corresponding quantity to a generalized momentum is

$$\Pi(x, t) = \frac{\partial\mathcal{L}}{\partial u_t}$$

- The Hamiltonian density is

$$\mathcal{H} = \frac{1}{2}\left(\rho u_t^2 + E_y u_x^2\right)$$

and the system's Hamiltonian is

$$H = S_0 \int \left(\Pi u_t - \mathcal{L}\right) \mathrm{d}x = S_0 \int \mathcal{H}\,\mathrm{d}x$$

- The Hamiltonian is conserved if $\frac{\mathrm{d}H}{\mathrm{d}t} = 0$. This occurs for

$$\frac{\mathrm{d}H}{\mathrm{d}t} = S_0 \int \frac{\mathrm{d}\mathcal{H}}{\mathrm{d}t}\,\mathrm{d}x = S \int \left(\rho u_t u_{tt} + E_y u_x u_{xx}\right)$$

With the wave equation $\rho u_{tt} = E u_{xx}$ and and a reverse product rule, this becomes

$$\frac{\mathrm{d}H}{\mathrm{d}t} = S_0 E_y \int \left(u_t u_{xx} + u_x u_{tx}\right) = S E_y \int \frac{\mathrm{d}}{\mathrm{d}x}[u_t u_x] = S E_y \big[u_t u_x\big]_0^l$$

where the rod's endpoints occur at $x = 0$ and $x = l$. If the rod is fixed on both ends, then

$$u_t(0,t) = u_t(l,t) = 0 \implies \frac{\mathrm{d}H}{\mathrm{d}t} = 0$$

and the system's energy is conserved.

3 Central Forces

3.1 One-Body Central Force Problem and Basic Equations of Orbit

What is the one-body central force problem? Discuss the problem's solution in both Newtonian and Lagrangian mechanics, making sure to explain the role of conserved quantities. What is the orbit equation? Discuss.

3.1.1 One-Body Central Force Problem: Lagrangian Approach

- Consider a body with mass m and position vector $\boldsymbol{r}$ in a central potential $V = V(r)$. The system's Lagrangian is

$$L = \frac{m\dot{\boldsymbol{r}}^2}{2} - V(r)$$

Because central forces act along the line connecting a body to the origin (i.e. $\boldsymbol{r} \parallel \boldsymbol{F}$), the body's angular momentum $\boldsymbol{L}$ is conserved. To see this, consider e.g.

$$\dot{\boldsymbol{L}} = \tau = \boldsymbol{r} \times \boldsymbol{F} = 0$$

Alternatively, the system's Lagrangian is invariant under infinitesimal rotations, so the system's angular momentum $\boldsymbol{L}$ is conserved by Noether's theorem. Since $\boldsymbol{L} = \boldsymbol{r} \times \boldsymbol{p}$ is perpendicular to $\boldsymbol{r}$ by definition and $\boldsymbol{L}$ is constant, the system's orbit must lie in a plane perpendicular to $\boldsymbol{L}$.

- Planar motion is best described in the polar coordinates (r, ϕ). In terms of r and ϕ the Lagrangian is

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

Since L is independent of ϕ, the quantity

$$p_\phi = \frac{\partial L}{\partial \dot{\phi}} = mr^2\dot{\phi} \equiv J$$

is conserved. J is the magnitude of angular momentum $|\boldsymbol{L}|$; I'm denoting it by J and not L to avoid confusion with the Lagrangian.

- The body's equation of motion comes from the Lagrange equation for r:

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{r}}\right) - \frac{\partial L}{\partial r} = m\ddot{r} - mr\dot{\phi}^2 + \frac{\partial V}{\partial r} = 0$$

We can then eliminate $\dot{\phi}$ by writing it in terms of the conserved quantity J

$$\dot{\phi}^2 = \frac{J^2}{m^2r^4} \implies m\ddot{r} = \frac{J^2}{mr^3} - \frac{\partial V}{\partial r} \equiv -\frac{\partial}{\partial r}V_{\text{eff}}(r)$$

where V_{eff} is the effective potential $V_{\text{eff}}(r) \equiv V(r) + \frac{J^2}{2mr^2}$. The equation of motion is

$$m\ddot{r} = -\frac{\partial V_{\text{eff}}}{\partial r} = \frac{J^2}{mr^3} - \frac{\partial V}{\partial r}$$

- Next, since the Lagrangian is not an explicit function of time (i.e. $\frac{\partial L}{\partial t} = 0$), the system's Hamiltonian is conserved, giving the conservation equation

$$E = \frac{m\dot{r}^2}{2} + V_{\text{eff}}(r)$$

This reduces the two-body central force problem to a one dimensional single body problem involving only r.

29

3.1.2   One-Body Central Force Problem: Newtonian Approach

- We start with a single particle of mass m with position vector $\boldsymbol{r}$ in a central force potential. The equation of motion for the one-particle system is

$$m\ddot{\boldsymbol{r}} = -\nabla V(r)$$

Angular momentum is conserved in a central potential. The proof is straightforward; we just have to show $\frac{\mathrm{d}\boldsymbol{L}}{\mathrm{d}t} = 0$. Here are two ways:

$$\frac{\mathrm{d}\boldsymbol{L}}{\mathrm{d}t} = m\dot{\boldsymbol{r}} \times \dot{\boldsymbol{r}} + \boldsymbol{r} \times m\ddot{\boldsymbol{r}} = 0 - \boldsymbol{r} \times \nabla V(r) = 0$$
$$\frac{\mathrm{d}\boldsymbol{L}}{\mathrm{d}t} = \boldsymbol{\tau} = \boldsymbol{r} \times \boldsymbol{F} = 0$$

The first equality holds because $\boldsymbol{r}$ and $\nabla V(r)$ are parallel for a central potential, and the second because $\boldsymbol{r}$ and $\boldsymbol{F}$ are parallel for a central force.

- Conservation of angular momentum means all motion in the central force problem takes place in a plane. Because $\boldsymbol{L} \cdot \boldsymbol{r} = 0$, the position of the particle is always perpendicular to $\boldsymbol{L}$. But because $\boldsymbol{L}$ is fixed and $\boldsymbol{r}$ is always perpendicular to $\boldsymbol{L}$, $\boldsymbol{r}$ must lie in a plane perpendicular to $\boldsymbol{L}$. Analogously, $\boldsymbol{L} \cdot \dot{\boldsymbol{r}} = 0$, so the particle's velocity $\dot{\boldsymbol{r}}$ also lies in the same plane perpendicular to $\boldsymbol{L}$ by the same argument.

- Planar dynamics is best analyzed with polar coordinates: we rotate the coordinate system so the z axis aligns with the conserved angular momentum $\boldsymbol{L}$ and define the usual polar coordinates

$$x = r\cos\phi \qquad \text{and} \qquad y = r\sin\phi$$

We then introduce the two unit vectors $\hat{\mathbf{r}}$ and $\hat{\phi}$; they point in direction of increasing r and ϕ, respectively. In Cartesian form, the unit vectors are written

$$\hat{\mathbf{r}} = \begin{bmatrix} \cos\phi \\ \sin\phi \end{bmatrix} \qquad \text{and} \qquad \hat{\phi} = \begin{bmatrix} -\sin\phi \\ \cos\phi \end{bmatrix}$$

The change in the basis $\{\hat{\mathbf{r}}, \hat{\phi}\}$ with respect to ϕ is

$$\frac{\mathrm{d}\hat{\mathbf{r}}}{\mathrm{d}\phi} = \begin{bmatrix} -\sin\phi \\ \cos\phi \end{bmatrix} = \hat{\phi} \qquad \text{and} \qquad \frac{\mathrm{d}\hat{\phi}}{\mathrm{d}\phi} = -\begin{bmatrix} \cos\phi \\ \sin\phi \end{bmatrix} = -\hat{\mathbf{r}}$$

In terms of the changing basis vectors $\{\hat{\mathbf{r}}, \hat{\phi}\}$ the position vector is simply

$$\boldsymbol{r} = r\hat{\mathbf{r}}$$

The velocity is found with the product rule for derivatives

$$\dot{\boldsymbol{r}} = \dot{r}\hat{\mathbf{r}} + r\frac{\mathrm{d}\hat{\mathbf{r}}}{\mathrm{d}t} = \dot{r}\hat{\mathbf{r}} + r\frac{\mathrm{d}\hat{\mathbf{r}}}{\mathrm{d}\phi}\frac{\mathrm{d}\phi}{\mathrm{d}t} = \dot{r}\hat{\mathbf{r}} + r\dot{\phi}\hat{\phi}$$

Finally, acceleration: a time differentiation of $\dot{\boldsymbol{r}}$ gives

$$\ddot{\boldsymbol{r}} = \ddot{r}\hat{\mathbf{r}} + \dot{r}\frac{\mathrm{d}\hat{\mathbf{r}}}{\mathrm{d}t} + \dot{r}\dot{\phi}\hat{\phi} + r\ddot{\phi}\hat{\phi} + r\dot{\phi}\frac{\mathrm{d}\hat{\phi}}{\mathrm{d}t}$$

Again, the chain rule and basic vector derivatives come in handy. Using the identities $\frac{\mathrm{d}\hat{\mathbf{r}}}{\mathrm{d}t} = \frac{\mathrm{d}\hat{\mathbf{r}}}{\mathrm{d}\phi}\frac{\mathrm{d}\phi}{\mathrm{d}t} = \dot{\phi}\hat{\phi}$ and $\frac{\mathrm{d}\hat{\phi}}{\mathrm{d}t} = \frac{\mathrm{d}\hat{\phi}}{\mathrm{d}\phi}\frac{\mathrm{d}\phi}{\mathrm{d}t} = -\dot{\phi}\hat{\mathbf{r}}$ we get

$$\ddot{\boldsymbol{r}} = (\ddot{r} - r\dot{\phi}^2)\hat{\mathbf{r}} + (r\ddot{\phi} + 2\dot{r}\dot{\phi})\hat{\phi}$$

30

- We return to the single-particle equation of motion for a central force

$$m\ddot{\boldsymbol{r}} = -\nabla V(r) = -\frac{\partial V}{\partial r}\hat{\mathbf{r}}$$

and substitute in the body's acceleration $\ddot{\boldsymbol{r}} = (\ddot{r} - r\dot{\phi}^2)\hat{\mathbf{r}} + (r\ddot{\phi} + 2\dot{r}\dot{\phi})\hat{\phi}$ to get

$$m(\ddot{r} - r\dot{\phi}^2)\hat{\mathbf{r}} + m(r\ddot{\phi} + 2\dot{r}\dot{\phi})\hat{\phi} = -\frac{\partial V}{\partial r}\hat{\mathbf{r}}$$

Notice the equation resolves into a radial component $(\ddot{r} - r\dot{\phi}^2)\hat{\mathbf{r}}$ and tangential component $(r\ddot{\phi} + 2\dot{r}\dot{\phi})\hat{\phi}$.

- Since there is no $\dot{\phi}$ term on the right hand side, the tangential equation of motion is

$$m(r\ddot{\phi} + 2\dot{r}\dot{\phi}) = 0$$

Rewriting $r\ddot{\phi} + 2\dot{r}\dot{\phi}$ as a time derivative shows

$$0 = m(r\ddot{\phi} + 2\dot{r}\dot{\phi}) = \frac{1}{r}\frac{\mathrm{d}}{\mathrm{d}t}(mr^2\dot{\phi}) \implies \frac{\mathrm{d}}{\mathrm{d}t}(mr^2\dot{\phi}) = 0$$

In other words, the quantity $mr^2\dot{\phi}$ is conserved in central force motion. The quantity $mr^2\dot{\phi}$ is magnitude of angular momentum $J \equiv |L|$; again, I'm writing J instead of L to avoid confusion with the Lagrangian.

To prove the equality we write $\boldsymbol{L}$ in terms of the basis vectors $\{\hat{\mathbf{r}}, \hat{\phi}\}$:

$$J \equiv |\boldsymbol{L}| = |m\boldsymbol{r} \times \dot{\boldsymbol{r}}| = \left|m(r\hat{\mathbf{r}}) \times (\dot{r}\hat{\mathbf{r}} + r\dot{\phi}\hat{\phi})\right| = mr^2\dot{\phi}\left|\hat{\mathbf{r}} \times \hat{\phi}\right| \implies J \equiv mr^2\dot{\phi}$$

- Returning to the radial component of the equation of motion and substituting in the constant of motion $\dot{\phi}^2 = \frac{L^2}{m^2 r^4}$ we get

$$m(\ddot{r} - r\dot{\phi}^2) = -\frac{\partial V}{\partial r} \implies m\ddot{r} = \frac{J^2}{mr^3} - \frac{\partial V}{\partial r} \equiv -\frac{\partial V_{\mathrm{eff}}}{\partial r}$$

where $V_{\mathrm{eff}}(r) = V(r) + \frac{J^2}{2mr^2}$ is the same effective potential as in the Lagrangian approach. This reduces the central force problem, originally in three dimensions, to a one-dimensional problem of radial motion.

- Next, we apply conservation of the system's total energy to get

$$
\begin{aligned}
E &= \frac{m\dot{\boldsymbol{r}}^2}{2} + V(r) = \frac{m}{2}\left(\dot{r}\hat{\mathbf{r}} + r\dot{\phi}\hat{\phi}\right)^2 + V(r) \\
&= \frac{m\dot{r}^2}{2} + \frac{mr^2\dot{\phi}^2}{2} + V(r) = \frac{m\dot{r}^2}{2} + \frac{J^2}{2mr^2} + V(r) \\
&= \frac{m\dot{r}}{2} + V_{\mathrm{eff}}(r)
\end{aligned}
$$

which is the same result as in the Lagrangian approach.

3.1.3   Example: Orbits with $V(r) = -\frac{k}{r}$ Classified by Energy

We'll consider the classic $\frac{1}{r}$ potential $V(r) = -\frac{k}{r}$. In this case the effective potential is

$$V_{\text{eff}} = -\frac{k}{r} + \frac{J^2}{2mr^2}$$

The minimum of the effective potential is:

$$r_{\text{min}} = \frac{J^2}{mk} \qquad \text{and} \qquad V_{\text{min}} = V(r_{\text{min}}) = -\frac{mk^2}{2J^2}$$

We can use these results to classify the possible forms of motion in a $\frac{1}{r}$ potential based on the system's energy E (it helps to graph the effective potential as a function or r). There are three options

1. If $E = V_{\text{min}}$ the particle rests at the minimum of the potential well V_{eff} and stays at the distance r_{min} from the origin. Although the particle has fixed radial position, it still moves with angular speed $\dot{\phi} = \frac{J}{mr_{\text{min}}^2}$: the particle moves in a circular orbit about the origin.

2. If $E \in (V_{\text{min}}, 0)$, the particle is trapped in a potential well and oscillates back and forth between the points of maximum energy. Since the particle also has angular speed $\dot{\phi} = \frac{J}{mr_{\text{min}}^2}$, the particle moves in a closed orbit in which radial distance r varies with time. For $V = -\frac{k}{r}$, this turns out to be an ellipse.

3. If $E > 0$, the particle has enough energy to overcome the potential well so the corresponding is unbounded. The particle approaches the origin from infinity, reaches a point of closest approach, and returns back out to infinity. For $V = -\frac{k}{r}$, the exact trajectory turns out to be a hyperbola.

3.1.4   The Orbit Equation

- *Orbits* are a general term for the trajectories of particles under central forces. It is straightforward to find how radial position $r(t)$ changes with time. Because of energy conservation we can write
$$E = \frac{m\dot{r}^2}{2} + V_{\text{eff}}(r)$$
and then invert the expression to give
$$\frac{\mathrm{d}r}{\mathrm{d}t} = \sqrt{\frac{2}{m}(E - V_{\text{eff}}(r))}$$

  This could then be integrated, e.g. numerically, to find an expression for $r(t)$.

- It turns out to be more useful to find radial position as a function of the polar angle ϕ; we then get a sense of the orbit by finding the shape of $r(\phi)$. To do this, we'll use the equation of motion
$$m\ddot{r} = -\frac{\partial V_{\text{eff}}}{\partial r} = \frac{J^2}{mr^3} - \frac{\partial V}{\partial r}$$
  Finding the expression for $r(\phi)$ begins with a trick: replace r with the new coordinate
$$u = \frac{1}{r}$$

We then use the chain rule and $\dot{\phi} = \frac{J}{mr^2}$ to find radial velocity $\dot{r}$

$$\frac{\mathrm{d}r}{\mathrm{d}t} = \frac{\mathrm{d}r}{\mathrm{d}\phi}\frac{\mathrm{d}\phi}{\mathrm{d}t} = \frac{\mathrm{d}r}{\mathrm{d}\phi}\frac{J}{mr^2} = -\frac{J}{m}\frac{\mathrm{d}u}{\mathrm{d}\phi}$$

where $\mathrm{d}u = -\frac{\mathrm{d}r}{r^2}$. Next, radial acceleration is

$$\frac{\mathrm{d}^2 r}{\mathrm{d}t^2} = -\frac{J}{m}\frac{\mathrm{d}}{\mathrm{d}t}\left[\frac{\mathrm{d}u}{\mathrm{d}\phi}\right] = -\frac{J}{m}\frac{\mathrm{d}^2 u}{\mathrm{d}\phi^2}\dot{\phi} = -\frac{J^2}{m^2 r^2}\frac{\mathrm{d}^2 u}{\mathrm{d}\phi^2} = -\frac{J^2}{m^2}\frac{\mathrm{d}^2 u}{\mathrm{d}\phi^2}u^2$$

where the $\dot{\phi}$ term comes from the chain rule and $u = u(\phi)$.

- In terms of u, the original orbit equation becomes

$$m\ddot{r} = -\frac{\mathrm{d}V}{\mathrm{d}r} + \frac{J^2}{mr^3} \iff \frac{\mathrm{d}^2 u}{\mathrm{d}\phi^2} + u = \frac{m}{J^2 u^2}\frac{\mathrm{d}V}{\mathrm{d}r}\bigg|_{r=\frac{1}{u}}$$

This is the *orbit equation*, which we want to solve for $u(\phi)$, which would then give us $r(\phi)$ via $u = \frac{1}{r}$. We could also extract the time dependence of r with $L = mr^2\dot{\phi}$.

To actually solve the orbit equation, we need to know the specific form of $V(r)$. The orbit equation is solved for the potential $V(r) = -\frac{k}{r}$ in the section on the Kepler problem.

3.2   Two-Body Central Force Problem

Discuss two-body central force problem. Explain how the problem is solved in both New-tonian and Lagrangian mechanics. Be sure to explain the role of conserved quantities in reducing the two-body problem to an equivalent one-body problem.

3.2.1   Lagrangian Approach

- As before, we have our two bodies with position vectors and masses $\boldsymbol{r}_1, m_1$ and $\boldsymbol{r}_2, m_2$, respectively. The center of mass $\boldsymbol{R}$ and reduced mass μ are

$$\boldsymbol{R} = \frac{m_1\boldsymbol{r}_1 + m_2\boldsymbol{r}_2}{m_1 + m_2} \qquad \text{and} \qquad \mu = \frac{m_1 m_2}{m_1 + m_2}$$

while the relative position vector $\boldsymbol{r}$ and its magnitude r are

$$\boldsymbol{r} = \boldsymbol{r}_1 - \boldsymbol{r}_2 \qquad \text{and} \qquad r = |\boldsymbol{r}|$$

The system's Lagrangian is

$$L = T - V = \frac{1}{2}\left(m_1\dot{\boldsymbol{r}}_1^2 + m_2\dot{\boldsymbol{r}}_2^2\right) - V(r)$$

- We start by writing the Lagrangian in terms of the center of mass $\boldsymbol{R}$ and relative position $\boldsymbol{r}$:

$$L = \frac{1}{2}(m_1 + m_2)\dot{\boldsymbol{R}}^2 + \frac{1}{2}\mu\dot{\boldsymbol{r}}^2 - V(r)$$

Note that the Lagrangian decomposes into a piece describing the center of mass $\boldsymbol{R}$ and a piece describing the relative separation $\boldsymbol{r}$.

Since there are no external forces acting on the system the motion of the center of mass is uniform, i.e. $\dot{\boldsymbol{R}}$. This means the $\frac{1}{2}(m_1 + m_2)\dot{\boldsymbol{R}}^2$ term is also constant and thus won't affect the system's dynamics.

- The system's Lagrangian is invariant under infinitesimal rotations, so the system's angular momentum $\boldsymbol{L}$ is conserved by Noether's theorem. Since $\boldsymbol{L} = \boldsymbol{r} \times \boldsymbol{p}$ is perpendicular to $\boldsymbol{r}$ by definition and $\boldsymbol{L}$ is constant, the system's orbit must lie in a plane perpendicular to $\boldsymbol{L}$.

  Planar motion is best suited to the polar coordinates (r, ϕ). In terms of r and ϕ the Lagrangian is

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

  Since L is independent of ϕ, the quantity

  $$\frac{\partial L}{\partial \dot{\phi}} = \mu r^2 \dot{\phi} \equiv J$$

  is conserved. J is the magnitude of angular momentum $|\boldsymbol{L}|$; I'm denoting it by J and not L to avoid confusion with the Lagrangian.

- We get the system's equations of motion from the Lagrange equations for the coordinate r:

  $$\frac{\mathrm{d}}{\mathrm{d}t} \left(\frac{\partial L}{\partial \dot{r}} \right) - \frac{\partial L}{\partial r} = \mu \ddot{r} - \mu r \dot{\phi}^2 + \frac{\partial V}{\partial r} = 0$$

  We can then eliminate $\dot{\phi}$ by writing it in terms of the conserved quantity J

  $$\dot{\phi}^2 = \frac{J^2}{\mu^2 r^4} \implies \mu \ddot{r} = \frac{J^2}{\mu r^3} - \frac{\partial V}{\partial r} \equiv -\frac{\partial}{\partial r} V_{\text{eff}}(r)$$

  where V_{eff} is the effective potential

  $$V_{\text{eff}}(r) = V(r) + \frac{J^2}{2\mu r^2}$$

- Next, since the Lagrangian is not an explicit function of time (i.e. $\frac{\partial L}{\partial t} = 0$), the system's Hamiltonian is conserved, given the energy conservation equation

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

  This reduces the two-body central force problem to a one dimensional single body problem involving only r. From here, we can reuse the results of the one-body central force problem essentially verbatim if just replace the single particle mass m with the reduced mass μ. Repeating the one-body procedure above leads to the two-body orbit equation

  $$\frac{\mathrm{d}^2 u}{\mathrm{d}\phi^2} + u = \frac{\mu}{J^2 u^2} \left. \frac{\mathrm{d}V}{\mathrm{d}r} \right|_{r = \frac{1}{u}}$$

  where the only difference is the replacement of m with the reduced mass μ.

3.2.2   Newtonian Approach

The two body problem is the name given to solving the dynamics of two mutually interacting particles. Conserved quantities make it possible to reduce the two body problem into an equivalent one-body problem. Here's how it works:

- Consider two bodies of mass m_1 and m_2 with position vectors $\boldsymbol{r}_1$ and $\boldsymbol{r}_2$. We denote the total mass by $M = m_1 + m_2$ and define the relative separation as

$$\boldsymbol{r} = \boldsymbol{r}_1 - \boldsymbol{r}_2$$

  Finally, we define the center of mass $\boldsymbol{R}$ with the equation

$$M\boldsymbol{R} = m_1\boldsymbol{r}_1 + m_2\boldsymbol{r}_2$$

- We begin simplifying the problem by writing

$$\boldsymbol{r}_1 = \boldsymbol{R} + \frac{m_2}{M}\boldsymbol{r} \qquad \text{and} \qquad \boldsymbol{r}_2 = \boldsymbol{R} - \frac{m_1}{M}\boldsymbol{r}$$

  Under the assumption that the two bodies are isolated, there is no external force on the system, which means the center of mass travels with constant velocity, i.e. $\ddot{\boldsymbol{R}} = 0$.

- Meanwhile, the relative motion of the two particles is governed by

$$\ddot{\boldsymbol{r}} = \ddot{\boldsymbol{r}}_1 - \ddot{\boldsymbol{r}}_2 = \frac{\boldsymbol{F}_{12}}{m_1} - \frac{\boldsymbol{F}_{21}}{m_2} = \left(\frac{m_1 + m_2}{m_1 m_2}\right)\boldsymbol{F}_{12}$$

  where $\boldsymbol{F}_{12}$ and $\boldsymbol{F}_{21}$ are the forces exerted by the second body on the first body and vice versa, respectively. The last equality uses Newton's third law, which tells us $\boldsymbol{F}_{12} = -\boldsymbol{F}_{21}$. It is customary (and efficient) to write the equation of motion for $\ddot{\boldsymbol{r}}$ in the condensed form

$$\mu\ddot{\boldsymbol{r}} = \boldsymbol{F}_{12}$$

  where we have defined the *reduced mass*

$$\mu = \frac{m_1 m_2}{m_1 + m_2}$$

  When one body is much more massive than the other (e.g. a planet orbiting a star, an electron orbiting a hydrogen atom, etc...), the reduced mass nearly equals the mass of the lighter body (e.g. $m_2 \gg m_1 \implies \mu \approx m_1$).

- The equation $\mu\ddot{\boldsymbol{r}} = \boldsymbol{F}_{12}$, turns two body problem into a one body problem involving a single position vector $\boldsymbol{r}$. We need to know the specific nature of the interaction $\boldsymbol{F}_{12}$ to procede further. If $\boldsymbol{F}_{12}$ is a *central force*, we can write

$$\boldsymbol{F} = -\nabla V(r) = -\frac{\mathrm{d}V}{\mathrm{d}r}\hat{\mathbf{r}}$$

  and reuse the results of the one-body central force problem essentially verbatim if just replace the single particle mass m with the reduced mass μ. Since the procedure from here is completely analogous to the one body problem, I won't repeat it here.

3.3   The Kepler Problem and Kepler's Laws

What is the Kepler problem? Discuss the relationship between the Kepler problem and the orbit equation. Solve the orbit equation, and discuss its implications for the orbits of planets. State and derive Kepler's laws of planetary motion.

3.3.1   The Kepler Problem

- The *Kepler problem* is just the name given to understanding planetary orbits about a star under the influence of gravity. The correct potential for this scenario is

$$V(r) = -\frac{GmM}{r}$$

where m and M are the mass of the planet and sun, respectively, while G is the universal gravitational constant. However, we will be a bit more general and work with the potential

$$V(r) = -\frac{k}{r}$$

with the tacit understanding that $k = GMm$. By using the general form, the results we derive will apply equally well to the electrostatic potential; we just have replace $k = \frac{-qQ}{4\pi\epsilon_0}$.

- The orbit equation from the two-body central force problem is

$$\frac{\mathrm{d}^2 u}{\mathrm{d}\phi^2} + u = \frac{\mu}{J^2 u^2} \frac{\mathrm{d}V}{\mathrm{d}r}\bigg|_{r=\frac{1}{u}}$$

where the reduced mass $\mu = \frac{mM}{m+M}$ replaces the single body mass m from the single body problem. For the potential $V = -\frac{k}{r}$, the orbit equation is simply

$$\frac{\mathrm{d}^2 u}{\mathrm{d}\phi^2} + u = \frac{\mu k}{J^2}$$

This is the equation for a harmonic oscillator with its center displaced by $\frac{\mu k}{J^2}$. We typically write the solution to this equation in the form $u = A\cos\phi + B\sin\phi + \frac{\mu k}{J^2}$, but for the orbit problem, it is much more useful to write

$$u(\phi) = A\cos(\phi - \phi_0) + \frac{\mu k}{J^2}$$

where the A and ϕ_0 are integration constants.

- We're almost there. At the orbit's point of closest approach, r is smallest and u is largest. From the our solution $u(\phi) = A\cos(\phi - \phi_0) + \frac{\mu k}{J^2}$, we see that u is largest when the cosine term equals one, meaning the point of closest approach occurs when $\phi - \phi_0 = 0$.

  This result provides a natural choice for our polar coordinate system: we orient our axes so the point of closest approach occurs at $\phi = 0$. To satisfy $\phi - \phi_0 = 0$ we then have the simple result $\phi_0 = 0$. We then have the solution

$$u(\phi) = A\cos\phi + \frac{\mu k}{J^2} \qquad \text{or} \qquad r(\phi) = \frac{1}{A\cos\phi + \frac{\mu k}{J^2}}$$

- By convention, we re-write the solution for $r(\phi)$ in the form

$$r(\phi) = \frac{r_0}{\epsilon\cos\phi + 1}$$

where $r_0 = \frac{J^2}{\mu k}$ and $\epsilon = \frac{AJ^2}{\mu k}$ (we just multiplied $r(\phi)$ above and below by $\frac{J^2}{\mu k}$).

  Note that r_0 is not an integration constant; it is uniquely determined by the fixed angular momentum J. The *eccentricity* ϵ is effectively the new integration constant.

3.3.2 Analyzing the Solutions

The solutions for $r(\phi)$ given by

$$r(\phi) = \frac{r_0}{\epsilon \cos \phi + 1}$$

are *conic sections* (e.g. ellipses, hyperbola, etc...). We can classify the orbits $r(\phi)$ by the eccentricity ϵ. There are a few choices:

1. For $\epsilon < 1$, the radial position is bounded in the interval

$$\frac{r_0}{r} \in [1 - \epsilon, 1 + \epsilon]$$

   since $\cos \phi$ falls between -1 and 1. This type of orbit is an ellipse. To double check, we can convert to Cartesian coordinates with the substitutions $x = r \cos \phi$, $y = \sin \phi$ and $r^2 = x^2 + y^2$. We have

$$r = \frac{r_0}{\epsilon \cos \phi + 1} \implies r^2 = (r_0 - \epsilon r \cos \phi)^2 \implies x^2 + y^2 = (r_0 - \epsilon x)^2$$

   A few more steps of algebra lead to the Cartesian equation of an ellipse

$$\frac{(x - x_0)^2}{a^2} + \frac{y^2}{b^2} = 1$$

   where

$$x_0 = -\frac{\epsilon r_0}{1 - \epsilon^2} \qquad a^2 = \frac{r_0^2}{(1 - \epsilon^2)^2} \qquad b^2 = \frac{r_0^2}{1 - \epsilon^2}$$

   The two axes of the ellipse have lengths a and b with $a > b$. The center of gravitational attraction (e.g. the sun) sits at the origin $r = 0$; the origin and ellipse's center are separated by

$$|x_0| = \frac{\epsilon r_0}{1 - \epsilon^2} = \epsilon a$$

   The origin of gravitational attraction is a *focus* of the ellipse; the other focus lies on the major axis an equal distance from the ellipse's center. Note that when $\epsilon = 0$, the focus and center coincide and the two axes have equal lengths $a = b$: this a circular orbit.

2. Orbits with $\epsilon > 1$ are hyperbolae. Repeating analogous algebraic steps to the elliptical orbit leads to the Cartesian coordinate equation

$$\frac{1}{a^2}(x + x_0)^2 - \frac{y^2}{b^2} = 1$$

   with a, b and x_0 defined as before.

   The hyperbola's asymptotes make angles $\cos \phi = -\frac{1}{\epsilon}$ with the x axis. This should make sense; from the polar orbit equation $r(\phi)$ we see that $r \to \infty$ when the denominator $\epsilon \cos \phi + 1$ equals zero; this occurs for $\cos \phi = -\frac{1}{\epsilon}$.

3. The special case $\epsilon = 1$ is particularly simple: it is a hyperbolic orbit. The Cartesian coordinate equation is

$$y^2 = r_0^2 - 2r_0 x$$

Classifying Solutions by Energy

- The solution of the orbit equation $r(\phi) = \frac{r_0}{\epsilon \cos \phi + 1}$ gives us an interesting new look at the orbit energy. Recall that

$$E = \frac{1}{2}\mu \dot{r}^2 + V_{\text{eff}}(r) = \frac{1}{2}\mu \dot{r}^2 - \frac{k}{r} + \frac{J^2}{2\mu r^2}$$

rewriting $\dot{r}$ with the chain rule and substituting $\dot{\phi} = \frac{J}{\mu r^2}$ gives

$$E = \frac{\mu}{2}\left(\frac{\mathrm{d}r}{\mathrm{d}\phi}\right)^2 \dot{\phi}^2 - \frac{k}{r} + \frac{J^2}{2mr^2}$$
$$= \frac{\mu}{2}\left(\frac{\mathrm{d}r}{\mathrm{d}\phi}\right)^2 \frac{J^2}{m^2 r^4} - \frac{k}{r} + \frac{J^2}{2\mu r^2}$$

- We can differentiate the solution $r(\phi) = \frac{r_0}{\epsilon \cos \phi + 1}$ to get

$$\frac{\mathrm{d}r}{\mathrm{d}\phi} = \frac{r_0 \epsilon \sin \phi}{(1 + \epsilon \cos \phi)^2}$$

Plugging this in to the energy expression and a few steps of rather satisfying algebra (using $\cos^2 \phi + \sin^2 \phi = 1$ and $r_0 = \frac{J^2}{\mu k}$) leads to the simple result

$$E = \frac{\mu k^2}{2J^2}(\epsilon^2 - 1)$$

Notice that all ϕ and r dependence has vanished from the expression for E. This makes sense: energy is a constant of motion.

- With the result $E = \frac{\mu k^2}{2J^2}(\epsilon^2 - 1)$ we now return the three cases of orbit classified by eccentricity to see that:

    1. Orbits with $\epsilon < 1$ have energy $E < 0$. Negative energy means a bounded orbit, in agreement with our earlier conclusion that orbits $\epsilon < 1$ are ellipses.

    2. Orbits with $\epsilon > 1$ have positive energy $E > 0$; these are unbounded orbits, in agreement with our earlier result that $\epsilon > 1$ leads to hyperbolic orbits.

    3. Finally, $\epsilon = 0$ leads to $E = -\frac{\mu k^2}{2J^2}$. We have seen this before: it is the minimum of the effective potential, and we know from earlier that orbits with minimum V_{eff} are circular, in agreement with our geometric interpretation that ellipses with $\epsilon = 0$ are circles.

3.3.3   Kepler's Laws of Planetary Motion

Kepler's three laws are obeyed by all planets in the solar system and were first published by Kepler in 1605.

- Kepler's first law states:

    Each planet moves in an ellipse, with the Sun at one focus.

    Without knowing, we have dedicated a good part of the past few sections to Kepler's first law when discussing the elliptical solution to the orbit equation (when $\epsilon < 1$).

- Kepler's second law is:

The line between the a planet and the Sun sweeps out equal areas in equal times at any point in the orbit.

The second law follows directly from conservation of angular momentum. From the equation of area in polar coordinates ($A = \frac{1}{2}\int r(\phi)^2 \, \mathrm{d}\phi$), we that in the time δt an orbiting planet sweeps out the area

$$\delta A = \frac{1}{2}r^2 \delta\phi$$

Differentiating both sides with respect to time and substituting in the conserved angular momentum $J = mr^2\dot{\phi} \implies \dot{\phi} = \frac{J}{\mu r^2}$ gives

$$\frac{\mathrm{d}A}{\mathrm{d}t} = \frac{1}{2}r^2\dot{\phi} = \frac{J}{2m} \equiv \frac{l}{2}$$

where $l \equiv \frac{J}{\mu} \equiv \frac{|\boldsymbol{L}|}{\mu}$ is the system's angular momentum per unit mass. Since l is a constant of motion, the quantity $\frac{\mathrm{d}A}{\mathrm{d}t}$ is also constant throughout the orbit. Because conservation of l holds for all central force motion, Kepler's second law would hold for any central potential $V = V(r)$, not just the gravitational potential $V = -\frac{GMm}{r}$.

- Kepler's third law states:

  The period of orbit is proportional the $\frac{3}{2}$ power of the radius, i.e. $T \sim r^{\frac{3}{2}}$.

We turn to the inverse square law of gravitation

$$F = \frac{GMm}{r^2}$$

There is a neat proof of the third law that follows simply from dimensional analysis. Kepler's third law is essentially a relation ship between a length and time describing the orbit (e.g. an average radius R and the period T). We recognize that the only parameter in the problem is GM, so the radius and period in Kepler's law must be related by GM.

The quantity GM has dimensions length$^3 \cdot$ time^{-2}, i.e. $[GM] = L^3 T^{-2}$. By dimensional analysis, Kepler's law relating R, T and GM must then be of the form

$$GM \sim \frac{R^3}{T^2} \quad \implies \quad T^2 \sim \frac{R^3}{GM} \quad \text{or} \quad T \sim R^{\frac{3}{2}}$$

- For an elliptical orbit we can be more precise. Using the area of an ellipse $A = \pi ab$ and our earlier relationships between a, b and r_0 we have

$$A = \pi ab = \pi a^2 \sqrt{1 - \epsilon^2} = \frac{\pi r_0^2}{(1 - \epsilon^2)^{3/2}}$$

Next, we use Kepler's second law: which tells us that area is swept out at the constant rate $\frac{\mathrm{d}A}{\mathrm{d}t} = \frac{l}{2} \implies \frac{A}{T} = \frac{l}{2}$ where $l = \frac{J}{\mu}$ is angular momentum per unit mass.

$$T = \frac{2A}{l} = \frac{2\pi r_0^2}{l(1 - \epsilon^2)^{3/2}} = \frac{2\pi}{\sqrt{GM}}\left(\frac{r_0}{1 - \epsilon^2}\right)^{3/2}$$

We see that T is indeed proportional to a radius quantity to the $\frac{3}{2}$ power. But what exactly does the quantity $\frac{r_0}{1 - \epsilon^2}$ represent? It turns out to have a very nice

interpretation. From the orbit equation $r(\phi)$ we see that the points of minimum and maximum approach occur at

$$r_{\min} = \frac{r_0}{(1 + \epsilon)} \qquad \text{and} \qquad r_{\max} = \frac{r_0}{(1 - \epsilon)}$$

respectively. The average of these quantities is none other than

$$R = \frac{1}{2}(r_{\min} + r_{\max}) = \frac{r_0}{(1 - \epsilon^2)}$$

R is the average radius of the orbit. In this interpretation, we have

$$T = \frac{2\pi}{\sqrt{GM}} R^{3/2}$$

proving Kepler's third law that the orbital period is proportional to the $\frac{3}{2}$ power of the (average) orbital radius.

3.4   The Laplace-Runge-Lenz Vector and Orbits

What is the Laplace-Runge-Lenz vector and how is it derived? Discuss the role of the LRL vector in solving the orbit equation.

3.4.1   Derivation

- Let's return to the two body central force problem, denoting the positions of the two bodies by $\boldsymbol{r}_1$ and $\boldsymbol{r}_2$ and defining the relative position vector $\boldsymbol{r}$ as

$$\boldsymbol{r} = \boldsymbol{r}_1 - \boldsymbol{r}_2$$

Because the force is assumed to be central, Newton's second law can be written

$$\mu\ddot{\boldsymbol{r}} = F(r)\hat{\mathbf{r}}$$

where $\hat{\mathbf{r}}$ is the unit vector in the direction of $\boldsymbol{r}$ and μ is the reduced mass. For the gravitational force $F(r) = \frac{-k}{r^2}$ with $k = GmM$, Newton's law becomes

$$\mu\ddot{\boldsymbol{r}} = \frac{-k}{r^2}\hat{\mathbf{r}}$$

- First, an intermediate step: we compute the time derivative $\dot{\hat{\mathbf{r}}}$. This seems odd, but it turns out super useful in the next step. The result is

$$\dot{\hat{\mathbf{r}}} = \frac{\dot{r}(\boldsymbol{r} \cdot \boldsymbol{r}) - (\boldsymbol{r} \cdot \dot{\boldsymbol{r}}) \cdot \boldsymbol{r}}{r^3} = \frac{r^2\dot{\boldsymbol{r}} - (\boldsymbol{r} \cdot \dot{\boldsymbol{r}})\boldsymbol{r}}{r^3}$$

Next, we take right-side cross product of both sides of Newton's law $\mu\ddot{\boldsymbol{r}} = \frac{-k}{r^2}\hat{\mathbf{r}}$ with the angular momentum $\boldsymbol{L} = \mu\boldsymbol{r} \times \dot{\boldsymbol{r}}$. This produces

$$\mu\ddot{\boldsymbol{r}} \times \boldsymbol{L} = \frac{-k}{r^2}\hat{\mathbf{r}} \times \boldsymbol{L} \implies \mu\ddot{\boldsymbol{r}} \times \boldsymbol{L} = -\frac{k\mu}{r^3}\boldsymbol{r} \times (\boldsymbol{r} \times \dot{\boldsymbol{r}})$$

We cancel masses, use the identity $\boldsymbol{a} \times (\boldsymbol{b} \times \boldsymbol{c}) = (\boldsymbol{a} \cdot \boldsymbol{c})\boldsymbol{b} - (\boldsymbol{a} \cdot \boldsymbol{b})\boldsymbol{c}$, and conveniently recognize the expression for $\dot{\hat{\mathbf{r}}}$:

$$\ddot{\boldsymbol{r}} \times \boldsymbol{L} = -\frac{k}{r^3}((\boldsymbol{r} \cdot \dot{\boldsymbol{r}})\boldsymbol{r} - r^2\dot{\boldsymbol{r}}) = +k\frac{(r^2\dot{\boldsymbol{r}} - (\boldsymbol{r} \cdot \dot{\boldsymbol{r}})\boldsymbol{r})}{r^3} = k\dot{\hat{\mathbf{r}}} \implies \ddot{\boldsymbol{r}} \times \boldsymbol{L} = k\dot{\hat{\mathbf{r}}}$$

- Because $\dot{\boldsymbol{L}} = 0$ ($\boldsymbol{L}$ is conserved in any central force problem), we can rewrite $\ddot{\boldsymbol{r}} \times \boldsymbol{L}$ as $\ddot{\boldsymbol{r}} \times \boldsymbol{L} = \frac{\mathrm{d}}{\mathrm{d}t}[\dot{\boldsymbol{r}} \times \boldsymbol{L}]$. Substituted into the result $\ddot{\boldsymbol{r}} \times \boldsymbol{L} = k\hat{\mathbf{r}}$, this gives us

$$\frac{\mathrm{d}}{\mathrm{d}t}[\dot{\boldsymbol{r}} \times \boldsymbol{L}] = k\dot{\hat{\mathbf{r}}}$$

Since k is a constant, we can integrate the equation with respect to time to get

$$\dot{\boldsymbol{r}} \times \boldsymbol{L} = k\hat{\mathbf{r}} + \boldsymbol{x}$$

where $\boldsymbol{x} \neq \boldsymbol{x}(t)$ is a time-independent vector resulting from the integration over time. (You could think of $\boldsymbol{x}$ as a vector integration constant).

- Finally, we divide the equation by k and rearrange to get

$$\boldsymbol{A} \equiv \frac{\boldsymbol{x}}{k} = \frac{\dot{\boldsymbol{r}} \times \boldsymbol{L}}{k} - \hat{\mathbf{r}}$$

where we have defined the Laplace-Runge-Lenz vector as $\boldsymbol{A} \equiv \frac{\boldsymbol{x}}{k}$. Because both $\boldsymbol{x}$ and k are time independent so is $\boldsymbol{A}$, i.e. the LRL vector is conserved for the $F \sim \frac{1}{r^2}$ central force problem.

3.4.2   The LRL Vector and the Orbit Equation

The LRL vector gives us a neat way of deriving the orbit equation for the Kepler problem of planetary motion.

- Starting with the definition of the LRL vector, we take the dot product of both sides with $\boldsymbol{r}$. With the help of the vector identity $(\boldsymbol{a} \times \boldsymbol{b}) \cdot \boldsymbol{c} = (\boldsymbol{c} \times \boldsymbol{a}) \cdot \boldsymbol{b}$ we get

$$\boldsymbol{A} \cdot \boldsymbol{r} = \frac{\dot{\boldsymbol{r}} \times \boldsymbol{L}}{k} \cdot \boldsymbol{r} - \hat{\mathbf{r}} \cdot \boldsymbol{r} = \frac{(\boldsymbol{r} \times \dot{\boldsymbol{r}}) \cdot \boldsymbol{L}}{k} - r = \frac{\boldsymbol{L} \cdot \boldsymbol{L}}{\mu k} - r = \frac{L^2}{k\mu} - r$$

- If we write $\boldsymbol{A} \cdot \boldsymbol{r} = Ar\cos\phi$ where ϕ is the angle between $\boldsymbol{A}$ and $\boldsymbol{r}$, and solve for r we get

$$r = \frac{L^2}{\mu k}\frac{1}{1 + A\cos\phi} \equiv \frac{r_0}{1 + A\cos\phi}$$

This should look familiar: it is the solution to the orbit equation for the Kepler problem with the perihelion set to $\phi = 0$. We had previously written the solution as

$$r(\phi) = \frac{r_0}{1 + \epsilon\cos\phi} \qquad \text{where} \qquad r_0 = \frac{L^2}{\mu k}$$

Comparing the two orbit equations reveals an awesome result: comparing the coefficient of the $\cos\phi$ term tells us

$$A = \epsilon$$

In other words, the eccentricity of the orbit is the magnitude of the LRL vector.

4 Motion of Rigid Bodies

4.1 Euler Angles and the Orientation of a Rigid Body

Explain how the Euler angles are used to encode the position and kinematics of a rotating rigid body.

4.1.1 Basic Description of Rigid Body

- Rigid bodies are objects where the distance between the points is fixed, i.e. $|\boldsymbol{r}_i - \boldsymbol{r}_j|$ is constant for all i, j. Rigid bodies have zero internal degrees of freedom because their constituent points are fixed.

- In general, a rigid body with a free center of mass has six degrees of freedom: three from rotation and three from translation. A rigid body with a *fixed* center of mass looses it translational freedom, so it has only three rotational degrees of freedom. The general motion of a free rigid body can be separated into a translation of the center of mass and rotation about a point, often chosen to be the center of mass.

- Conventionally rigid bodies are analyzed in a fixed (inertial) lab frame $\{\tilde{\boldsymbol{e}}_a\}$ and a body-fixed frame $\{\boldsymbol{e}_a\} = \{\boldsymbol{e}_a\}(t)$ that moves with the body. Both frames have orthonormal axes: $\tilde{\boldsymbol{e}}_a \cdot \tilde{\boldsymbol{e}}_b = \delta_{ab}$ and $\boldsymbol{e}_a(t) \cdot \boldsymbol{e}_b(t) = \delta_{ab}$.

- For all times t, there exists a *unique* rotation matrix $\mathbf{R}(t)$ for which $\boldsymbol{e}_a(t) = \mathbf{R}_{ab}(t)\tilde{\boldsymbol{e}}_b$. In other words, it is always possible to rotate either the lab frame or body-fixed frame in such a way that the directions of the axes $\{\tilde{\boldsymbol{e}}_a\}$ and $\{\boldsymbol{e}_a\}$ coincide, and the corresponding rotation matrix $\mathbf{R}$ is unique.

  As the rigid body rotates, we can thus describe its orientation with the 3×3 time-dependent rotation matrix $\mathbf{R}(t)$. As for any rotation matrix $\mathbf{R}$, is orthogonal, $\mathbf{R}^T\mathbf{R} = \mathbf{I}$ and $\det \mathbf{R} = 1$. In general, a 3×3 matrix has 9 degrees of freedom, but the orthogonality condition $\mathbf{R}^T\mathbf{R} = \mathbf{I}$ imposes 6 constraints, leading to 3 degrees of rotational freedom. We thus need three coordinates to parameterize a rigid body's rotation.

4.1.2 Euler Angles and Euler's Rotation Theorem

- Informally, Euler angles are just coordinates well-suited to describing the dynamics of a rigid body. More exactly, the Euler angles are parametrization of the rotation matrix $\mathbf{R}$ from the inertial lab frame $\{\tilde{\boldsymbol{e}}_a\}$ to the body-fixed frame $\{\boldsymbol{e}_a\}$. The angles encode how the body-fixed system of principal axes of a rigid body is oriented relative to an inertial lab frame.

- The Euler angles rest on the theoretical framework of *Euler's rotation theorem*, which tells us:

  > An arbitrary rotation in three dimensions can be expressed as the product of 3 successive rotations about 3 (in general) different axes.

- Calling on Euler's theorem, we can get from the inertial lab system $\{\tilde{\boldsymbol{e}}_a\}$ to the body-fixed system $\{\boldsymbol{e}_a\}$ in three steps, using the rotations $\mathrm{R}_3(\phi), \mathrm{R}_1(\theta)$ and $\mathrm{R}_3(\psi)$

$$\{\tilde{\boldsymbol{e}}_a\} \overset{\mathrm{R}_3(\phi)}{\to} \{\boldsymbol{e}_a'\} \overset{\mathrm{R}_1(\theta)}{\to} \{\boldsymbol{e}_a''\} \overset{\mathrm{R}_3(\psi)}{\to} \{\boldsymbol{e}_a\}$$

where $\mathrm{R}_a(\phi)$ means a rotation about the ath basis vector by the angle ϕ, e.g. $\mathrm{R}_3(\phi)$ means a rotation about $\tilde{\boldsymbol{e}}_3$ by the angle ϕ. The angles of rotation ϕ, θ and ψ are called the *Euler angles*. In detail, the three steps are:

1. Rotate about $\tilde{\boldsymbol{e}}_3$ by the angle ϕ. ϕ is called the precession angle. We have

$$\boldsymbol{e}'_a = \mathrm{R}_3(\phi)_{ab}\tilde{\boldsymbol{e}}_b \qquad \mathbf{R}_3(\phi) = \begin{bmatrix} \cos\phi & \sin\phi & 0 \\ -\sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

2. Rotate about $\boldsymbol{e}'_1$ by the angle θ. θ is called the nutation angle.

$$\boldsymbol{e}''_a = \mathrm{R}_1(\theta)_{ab}\boldsymbol{e}'_b \qquad \mathbf{R}_1(\theta) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & \sin\theta \\ 0 & -\sin\theta & \cos\theta \end{bmatrix}$$

3. Rotate about $\boldsymbol{e}''_3$ by angle ψ to reach the body-fixed frame $\{\boldsymbol{e}_a\}$. ψ corresponds to rotation of the body about its figure axis.

$$\boldsymbol{e}_a = \mathrm{R}_3(\psi)_{ab}\boldsymbol{e}''_b \qquad \mathbf{R}_3(\psi) = \begin{bmatrix} \cos\psi & \sin\psi & 0 \\ -\sin\psi & \cos\psi & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

- The angles ϕ, θ, ψ are the Euler angles. The complete rotation matrix transforming from the lab frame to the body-fixed frame is

$$\mathbf{R}(\phi,\theta,\psi) = \mathbf{R}_3(\psi)\mathbf{R}_1(\theta)\mathbf{R}_3(\phi) \qquad \text{and} \qquad \boldsymbol{e}_a = \mathrm{R}_{ab}(\phi,\theta,\psi)\tilde{\boldsymbol{e}}_b$$

In full, the matrix reads

$$\mathbf{R}(\phi,\theta,\psi) = \begin{bmatrix} \cos\psi\cos\phi - \cos\theta\sin\phi\sin\psi & \sin\phi\cos\psi + \cos\theta\sin\psi\cos\phi & \sin\theta\sin\psi \\ -\cos\phi\sin\psi - \cos\theta\cos\psi\sin\phi & -\sin\psi\sin\phi + \cos\theta\cos\psi\cos\phi & \sin\theta\cos\psi \\ \sin\theta\sin\phi & -\sin\theta\cos\phi & \cos\theta \end{bmatrix}$$

Note that the rotations as defined above are *passive rotations*: they rotate the coordinate axes of the lab frame $\{\tilde{\boldsymbol{e}}_a\}$ to coincide with the lab frame.

- In terms of the body-fixed frame $\{\boldsymbol{e}_a\}$, the basis vectors $\tilde{\boldsymbol{e}}_3$, $\boldsymbol{e}'_1$ and $\boldsymbol{e}_3$ are

$$\tilde{\boldsymbol{e}}_3 = \mathbf{R}_3(\psi)\mathbf{R}_1(\theta)\mathbf{R}_3(\phi)\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \sin\theta\sin\psi\boldsymbol{e}_1 + \sin\theta\cos\psi\boldsymbol{e}_2 + \cos\theta\boldsymbol{e}_3$$

$$\boldsymbol{e}'_1 = \mathbf{R}_3(\psi)\mathbf{R}_1(\theta)\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \cos\psi\boldsymbol{e}_1 - \sin\psi\boldsymbol{e}_2$$

$$\boldsymbol{e}_3 = \boldsymbol{e}_3 \quad \text{(obviously)}$$

- By the construction of the Euler angles, the angular velocity can be written

$$\boldsymbol{\omega} = \dot{\phi}\tilde{\boldsymbol{e}}_3 + \dot{\theta}\boldsymbol{e}'_1 + \dot{\psi}\boldsymbol{e}_3$$

where the basis vectors are the axes of rotation in each step of the Euler angles: $\tilde{\boldsymbol{e}}_3$ is the lab-frame z axis, $\boldsymbol{e}'_1$ is the intermediate x' axis after the first rotation, and $\boldsymbol{e}_3$ is the object's figure axis.

- In terms of the body-fixed frame $\{e_a\}$, the basis vectors $\tilde{e}_3, e_1'$ and e_3 read

$$\tilde{e}_3 = \sin\theta\sin\psi e_1 + \sin\theta\cos\psi e_2 + \cos\theta e_3$$
$$e_1' = \cos\psi e_1 - \sin\psi e_2$$
$$e_3 = e_3$$

- Plugging the basis vectors into $\boldsymbol{\omega} = \dot{\phi}\tilde{e}_3 + \dot{\theta}e_1' + \dot{\psi}e_3$ and grouping like terms produces

$$\boldsymbol{\omega} = (\dot{\phi}\sin\theta\sin\psi + \dot{\theta}\cos\psi)e_1 + (\dot{\phi}\sin\theta\cos\psi - \dot{\theta}\sin\psi)e_2 + (\dot{\phi}\cos\theta + \dot{\psi})e_3$$

the desired expression for angular velocity in the body-fixed system.

4.2   Kinematics of a Rigid Body With One Fixed Point

Provide a thorough treatment of the kinematics of a rigid body with on fixed point. Be sure to mention angular velocity, the inertia tensor, angular momentum, and the Euler equations.

4.2.1   Angular Velocity, Formal Approach

This section relies heavily on the Einstein summation convention.

- Any point in the body $\boldsymbol{r}(t)$ can be expanded in either the body frame or space frame:

$$\boldsymbol{r}(t) = r_a \boldsymbol{e}_a(t) = \tilde{r}_a(t)\tilde{e}_a$$

where $\tilde{r}_b(t) = r_a \mathrm{R}_{ab}(t)$ and $\mathbf{R}(t)$ is the rotation matrix transforming from the space frame to the body frame. In the body-fixed frame, the distances r_a are constant and in the lab frame, the basis vectors $\tilde{e}_a$ are constant.

- Taking the time derivative of $\boldsymbol{r}(t)$ produces

$$\dot{\boldsymbol{r}}(t) = \tilde{e}_a \frac{\mathrm{d}\tilde{r}_a}{\mathrm{d}t} \qquad \text{(lab frame)}$$
$$= r_a \frac{\mathrm{d}\boldsymbol{e}_a}{\mathrm{d}t} \qquad \text{(body-fixed frame)}$$
$$= r_a \frac{\mathrm{d}\mathrm{R}_{ab}}{\mathrm{d}t}\tilde{e}_b \quad \left(\text{using } \boldsymbol{e}_a = \mathrm{R}_{ab}\,\tilde{e}_b\right)$$

- Using the transformation $\boldsymbol{e}_b = \mathrm{R}_{ab}\,\tilde{e}_a$ and the fact that the lab basis vectors $\tilde{e}_a$ are constant, the time derivative of the body-fixed basis vectors $\{\boldsymbol{e}_a(t)\}$ is

$$\frac{\mathrm{d}\boldsymbol{e}_a}{\mathrm{d}t} = \frac{\mathrm{d}}{\mathrm{d}t}[\mathrm{R}_{ab}\,\tilde{e}_b] = \frac{\mathrm{d}\mathrm{R}_{ab}}{\mathrm{d}t}\tilde{e}_b$$

Next, using the transformation $\tilde{e}_b = \mathrm{R}_{bc}^{-1}\,\boldsymbol{e}_c$, and $\mathbf{R}$'s orthogonality $\mathrm{R}_{ij}^{-1} = \mathrm{R}_{ji}$ we finish the calculation with

$$\frac{\mathrm{d}\boldsymbol{e}_a}{\mathrm{d}t} = \cdots = \frac{\mathrm{d}\mathrm{R}_{ab}}{\mathrm{d}t}(\mathrm{R}_{bc}^{-1}\,\boldsymbol{e}_c) = \frac{\mathrm{d}\mathrm{R}_{ab}}{\mathrm{d}t}(\mathrm{R}_{cb}\,\boldsymbol{e}_c) \equiv \Omega_{ac}\,\boldsymbol{e}_c$$

where we've defined $\Omega_{ac} \equiv \frac{\mathrm{d}\mathrm{R}_{ab}}{\mathrm{d}t}\mathrm{R}_{cb}$, the elements of a 3×3 matrix $\boldsymbol{\Omega}$.

The matrix $\boldsymbol{\Omega}$ defined above is antisymmetric, i.e. $\Omega_{ac} = -\Omega_{ca}$. To show this, we differentiate the identity $\mathrm{R}_{ab}\mathrm{R}_{cb} = \delta_{ac}$ and recognize the definition of Ω_{ac}

$$\mathrm{R}_{ab}\mathrm{R}_{cb} = \delta_{ac} \implies \frac{\mathrm{d}\mathrm{R}_{ab}}{\mathrm{d}t}\mathrm{R}_{cb} + \mathrm{R}_{ab}\frac{\mathrm{d}\mathrm{R}_{cb}}{\mathrm{d}t} = 0 \implies \Omega_{ac} + \Omega_{ca} = 0$$

- Because $\boldsymbol{\Omega}$ is antisymmetric (its diagonal terms are necessarily zero and its off-diagonal terms are related by $\Omega_{ab} = -\Omega_{ba}$) it has only 3 independent components. To be concrete, $\boldsymbol{\Omega}$ looks something like this:

$$\boldsymbol{\Omega} = \begin{bmatrix} 0 & \Omega_{12} & \Omega_{13} \\ \Omega_{21} & 0 & \Omega_{23} \\ \Omega_{31} & \Omega_{32} & 0 \end{bmatrix} = \begin{bmatrix} 0 & \Omega_{12} & -\Omega_{31} \\ -\Omega_{12} & 0 & \Omega_{23} \\ \Omega_{31} & -\Omega_{23} & 0 \end{bmatrix}$$

Since $\boldsymbol{\Omega}$ has only 3 independent components, we can define an equivalent three-component, one-index object (i.e. a vector) $\boldsymbol{\omega}$ with the formula

$$\omega_a = \frac{1}{2}\epsilon_{abc}\,\Omega_{bc} \qquad a, b, c = 1, 2, 3$$

By components, we get $\omega_1 = \Omega_{23}$, $\omega_2 = \Omega_{31}$ and $\omega_3 = \Omega_{12}$ (try it yourself!).

- The ω_a form the components of a vector in the body-fixed frame $\{\boldsymbol{e}_a\}$: $\boldsymbol{\omega} = \omega_a\boldsymbol{e}_a$. In terms of $\boldsymbol{\omega}$, the time derivative of the body frame $\{\boldsymbol{e}_a\}$ is

$$\dot{\boldsymbol{e}}_a(t) = \Omega_{ac}\,\boldsymbol{e}_c = -\epsilon_{abc}\omega_b\boldsymbol{e}_c = w_b\boldsymbol{e}_b \times \boldsymbol{e}_a = \boldsymbol{\omega} \times \boldsymbol{e}_a$$

where $\boldsymbol{e}_a \times \boldsymbol{e}_b = \epsilon_{abc}\boldsymbol{e}_c$. This is the final expression for the time derivative of the body-fixed frame. I'll stress again that its components ω_a are measured with respect to the body-fixed frame $\{\boldsymbol{e}_a\}$. Angular velocity comes up everywhere in rigid body mechanics, we'll se much more of it in the next sections.

- Informal interpretation: the angular velocity encodes the speed and direction of the body's rotation. To see this, consider a point $\boldsymbol{r}$ in the body, and rotate the point by the infinitesimal angle $\mathrm{d}\phi$ about the fixed axis $\hat{\mathbf{n}}$, where the angle between $\boldsymbol{r}$ and $\hat{\mathbf{n}}$ is θ. The vector $\boldsymbol{r}$ changes in magnitude by $|\mathrm{d}\boldsymbol{r}| = |\boldsymbol{r}|\sin\theta\,\mathrm{d}\phi$. Because the body is rigid, $\mathrm{d}\boldsymbol{r} \perp \boldsymbol{r}$, so $\mathrm{d}\boldsymbol{r} = \mathrm{d}\boldsymbol{\phi} \times \boldsymbol{r}$, where $\mathrm{d}\boldsymbol{\phi} = \mathrm{d}\phi\hat{\mathbf{n}}$. "Dividing" $\mathrm{d}\boldsymbol{r} = \mathrm{d}\boldsymbol{\phi} \times \boldsymbol{r}$ by $\mathrm{d}t$ produces

$$\frac{\mathrm{d}\boldsymbol{r}}{\mathrm{d}t} = \frac{\mathrm{d}\boldsymbol{\phi}}{\mathrm{d}t} \times \boldsymbol{r} = \boldsymbol{\omega} \times \boldsymbol{r}$$

Thus, $\boldsymbol{\omega}$ is the body's instantaneous angular velocity about the rotation axis $\hat{\mathbf{n}}$. In general, both $\hat{\mathbf{n}}$ and $\boldsymbol{\omega}$ change over time.

4.2.2   Inertia Tensor

- The inertia tensor arises naturally from computing the kinetic energy of a fixed rotating body:

$$T = \frac{1}{2}\sum_i m_i\dot{\boldsymbol{r}}_i \cdot \dot{\boldsymbol{r}}_i = \frac{1}{2}\sum_i m_i(\boldsymbol{\omega} \times \boldsymbol{r}_i) \cdot (\boldsymbol{\omega} \times \boldsymbol{r}_i)$$

$$= \frac{1}{2}\sum_i m_i\left[(\boldsymbol{\omega} \cdot \boldsymbol{\omega})(\boldsymbol{r}_i \cdot \boldsymbol{r}_i) - (\boldsymbol{\omega} \cdot \boldsymbol{r}_i)^2\right]$$

where the last equality uses Lagrange's vector identity $(\boldsymbol{a} \times \boldsymbol{b})^2 = \boldsymbol{a}^2\boldsymbol{b}^2 - (\boldsymbol{a} \cdot \boldsymbol{b})^2$. The equality can be written

$$T = \frac{1}{2}\omega_a\,\mathrm{I}_{ab}\,\omega_b = \frac{1}{2}\boldsymbol{\omega} \cdot \mathbf{I}\boldsymbol{\omega}$$

where $\mathbf{I}$ is the *inertia tensor*. Its components, *measured in the body-fixed frame*, are

$$\mathrm{I}_{ab} = \sum_i m_i \left((\boldsymbol{r}_i \cdot \boldsymbol{r}_i)\delta_{ab} - r_{i_a} r_{i_b} \right)$$

A few important notes:

- By construction $\mathrm{I}_{ab} = \mathrm{I}_{ba}$, so $\mathbf{I}$ is symmetric.
- Since $\mathbf{I}$ is measured in the body-fixed frame, its components are independent of time, i.e. $\dot{\mathbf{I}} = 0$
- A body's inertia tensor is always measured with respect to a particular point in space, so the inertia tensor is in general different when measured with respect to different points in space.

For a continuous mass distribution with density $\rho = \rho(\boldsymbol{r})$, the inertia tensor is

$$\mathbf{I} = \int \mathrm{d}^3\boldsymbol{r}\rho(\boldsymbol{r}) \begin{bmatrix} y^2 + z^2 & -xy & -xz \\ -xy & x^2 + z^2 & -yz \\ -xz & -yz & x^2 + y^2 \end{bmatrix}$$

- Because $\mathbf{I}$ is real and symmetric, it can be diagonalized, i.e. there exists an orthogonal matrix $\mathbf{Q}$ such that $\mathbf{I}' = \mathbf{Q}\mathbf{I}\mathbf{Q}^T$ is diagonal.

Equivalently, we can rotate the orientation of the body-fixed frame $\{\boldsymbol{e}_a\}$ to coincide with $\mathbf{I}$'s eigenvectors, which are given by $\{\mathbf{Q}\boldsymbol{e}_a\}$. In the frame $\{\mathbf{Q}\boldsymbol{e}_a\}$ the inertia tensor $\mathbf{I}$ is diagonal.

The frame in which $\mathbf{I}$ is diagonal is called the *frame of principle axes*. In this basis corresponding to the principle axes frame, $\mathbf{I}$ takes the form

$$\mathbf{I} = \begin{bmatrix} I_1 & 0 & 0 \\ 0 & I_2 & 0 \\ 0 & 0 & I_3 \end{bmatrix}$$

$\mathbf{I}$'s eigenvalues $I_{1,2,3}$ are called the *principle moments of inertia*. The frame of principle axes is the canonical choice for the body-fixed frame, and nearly all analyses will choose to align the body fixed frame with the principle axes. By convention, the origin is placed at the body's center of mass.

4.2.3   Angular Momentum

- Angular momentum, like angular velocity and the inertia tensor, is calculated with respect to a given point. In terms of the inertia tensor, using the vector triple product, we have:

$$\begin{aligned} \boldsymbol{L} &= \sum_i m_i \boldsymbol{r}_i \times \dot{\boldsymbol{r}}_i = \sum_i m_i \boldsymbol{r}_i \times (\boldsymbol{\omega} \times \boldsymbol{r}_i) \\ &= \sum_i m_i (r_i^2 \boldsymbol{\omega} - (\boldsymbol{r}_i \cdot \boldsymbol{\omega})\boldsymbol{r}_i) \\ &= \mathbf{I}\boldsymbol{\omega} \end{aligned}$$

- There is an important concept hiding in the result $\boldsymbol{L} = \mathbf{I}\boldsymbol{\omega}$: in general angular velocity $\boldsymbol{\omega}$ and angular momentum $\boldsymbol{L}$ point in different directions.

46

There are probably more elegant ways of showing this, but this way worked for me. By components, $\boldsymbol{L} = \mathbf{I}\boldsymbol{\omega}$ reads

$$
\begin{bmatrix} L_1 \\ L_2 \\ L_3 \end{bmatrix} = \begin{bmatrix} I_{11}\omega_1 + I_{12}\omega_2 + I_{13}\omega_3 \\ I_{21}\omega_1 + I_{22}\omega_2 + I_{23}\omega_3 \\ I_{31}\omega_1 + I_{32}\omega_2 + I_{33}\omega_3 \end{bmatrix} \implies L_a = I_{ab}\omega_b
$$

In other words, each component L_a is a linear combination of *all* of $\boldsymbol{\omega}$'s components, which means that in general $\boldsymbol{L}$ and $\boldsymbol{\omega}$ different directions; $\boldsymbol{L}$ and $\boldsymbol{\omega}$ coincide only for rotation about a principle axis of inertia.

4.2.4   Rigid Body Motion as Translation and Rotation

- Everything we've said so far applies to a body rotating about a fixed point; i.e. the point of rotation has not moved through space.

  More generally, the motion of a rigid body can always be written as a translation through space plus a rotation about a point in space. In theory, the point of rotation could be anywhere inside or outside the body, but is conventionally chosen to be the center of mass.

- We can decompose the position $\boldsymbol{r}_i$ of a point in the body into

$$
\boldsymbol{r}_i(t) = \boldsymbol{R}(t) + \Delta\boldsymbol{r}_i(t)
$$

  where $\boldsymbol{R}(t)$ is the center of mass and $\Delta\boldsymbol{r}_i(t)$ is the position of the point relative to the center of mass. With the decomposition $\boldsymbol{r}_i = \boldsymbol{R} + \Delta\boldsymbol{r}_i$, the kinetic energy of a rotating body decomposes into a translation of the center of mass plus rotation about the center of mass.

$$
\begin{aligned}
T &= \frac{1}{2}\sum_i m_i(\dot{\boldsymbol{r}}_i \cdot \dot{\boldsymbol{r}}_i) = \frac{1}{2}\sum_i m_i \left[(\dot{\boldsymbol{R}} + \Delta\dot{\boldsymbol{r}}_i) \cdot (\dot{\boldsymbol{R}} + \Delta\dot{\boldsymbol{r}}_i) \right] \\
&= \frac{1}{2}\sum_i m_i \left(\dot{\boldsymbol{R}}^2 + 2\{\dot{\boldsymbol{R}} \cdot (\boldsymbol{\omega} \times \Delta\boldsymbol{r}_i)\} + (\boldsymbol{\omega} \times \Delta\boldsymbol{r}_i)^2 \right) \\
&= \frac{1}{2}\sum_i m_i \dot{\boldsymbol{R}}^2 + \frac{1}{2}\omega_a \, \mathrm{I}_{ab} \, \omega_b = \frac{1}{2}M\dot{\boldsymbol{R}}^2 + \frac{1}{2}\boldsymbol{\omega} \cdot \mathbf{I}\boldsymbol{\omega}
\end{aligned}
$$

  where the term in curly brackets $\{\cdots\}$ is linear in $\Delta\boldsymbol{r}_i$ vanishes by definition of center of mass $\sum_i m_i\boldsymbol{r}_i = 0$ and $\mathbf{I}$ is measured with respect to the center of mass.

  Since the kinetic energy separates into translational and rotational terms, it makes sense to analyze rotation separately, at least for a free body in the absence of a potential energy.

4.2.5   Euler Equations

For now, we'll focus only on the rotational (i.e. not translational) motion of a rigid body. We'll start with a free body, then generalize to allow for the presence of an external torque.

- The angular momentum of any free body must be conserved, so $\dot{\boldsymbol{L}} = \boldsymbol{0}$, which allows us to define the body's equation of motion. Using the body-fixed expansion of angular momentum $\boldsymbol{L} = L_a\boldsymbol{e}_a$ the equations of motion are

$$
\begin{aligned}
\boldsymbol{0} \equiv \frac{\mathrm{d}\boldsymbol{L}}{\mathrm{d}t} &= \frac{\mathrm{d}L_a}{\mathrm{d}t}\boldsymbol{e}_a + L_a \frac{\mathrm{d}\boldsymbol{e}_a}{\mathrm{d}t} \\
&= \frac{\mathrm{d}L_a}{\mathrm{d}t}\boldsymbol{e}_a + L_a\boldsymbol{\omega} \times \boldsymbol{e}_a
\end{aligned}
$$

- In principle, the body-fixed frame could be any frame rotating with the body. But the equations of motion simplify considerably if we choose the body-fixed frame to align with the system of principle axes in which $\mathbf{I}$ is diagonal.

  If we align the body-fixed frame with the principle axes, we get the simple expression $L_a = \mathrm{I}_a\,\omega_a$ with no summation implied over a, e.g. $L_1 = I_1\omega_1$.

  The equations of motion become

  $$I_1\dot{\omega}_1 + (I_3 - I_2)\omega_2\omega_3 = 0$$
  $$I_2\dot{\omega}_2 + (I_1 - I_3)\omega_3\omega_1 = 0$$
  $$I_3\dot{\omega}_3 + (I_2 - I_1)\omega_1\omega_2 = 0$$

  The equations are called the *Euler equations*. They are the equations of motion for free rigid body rotating about its center of mass; you use them to calculate the body's angular velocity $\boldsymbol{w}$ in the body-fixed frame. Since they really on $\mathbf{I}$ being diagonal, they apply *only* if the the body-fixed frame aligns with the system of principle of axes measured about the point of rotation. If the principle axes and body-fixed frame aren't aligned, you have to use the more general equation

  $$\mathbf{0} = \frac{\mathrm{d}L_a}{\mathrm{d}t}\boldsymbol{e}_a + L_a\boldsymbol{\omega} \times \boldsymbol{e}_a$$

  Here's a quick derivation of the Euler equations using column vectors. Note that $\frac{\mathrm{d}L_a}{\mathrm{d}t} = \frac{\mathrm{d}}{\mathrm{d}t}[I_a\omega_a] = I_a\dot{\omega}_a$ because the I_a are constant in the body frame.

  $$
  \begin{aligned}
  \mathbf{0} &= \frac{\mathrm{d}L_a}{\mathrm{d}t}\boldsymbol{e}_a + L_a\boldsymbol{\omega} \times \boldsymbol{e}_a = I_a\dot{\omega}_a\boldsymbol{e}_a + I_a\omega_a\boldsymbol{\omega} \times \boldsymbol{e}_a \\
  &= \begin{bmatrix} I_1\dot{\omega}_1 \\ I_2\dot{\omega}_2 \\ I_3\dot{\omega}_3 \end{bmatrix} + I_1\omega_1\begin{bmatrix} 0 \\ \omega_3 \\ -\omega_2 \end{bmatrix} + I_2\omega_2\begin{bmatrix} -\omega_3 \\ 0 \\ \omega_1 \end{bmatrix} + I_3\omega_3\begin{bmatrix} \omega_2 \\ -\omega_1 \\ 0 \end{bmatrix} \\
  &= \begin{bmatrix} I_1\dot{\omega}_1 + (I_3 - I_2)\omega_2\omega_3 \\ I_2\dot{\omega}_2 + (I_1 - I_3)\omega_3\omega_1 \\ I_3\dot{\omega}_3 + (I_2 - I_1)\omega_1\omega_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}
  \end{aligned}
  $$

- We can also generalize Euler's equations to apply to a body acted on by an external torque $\boldsymbol{M}$. Instead of starting with $\dot{\boldsymbol{L}} = \mathbf{0}$ we start from $\dot{\boldsymbol{L}} = \boldsymbol{M}$, expanding the vectors in the body-fixed frame and following an analogous derivation. The Euler equations with an external torque are

  $$I_1\dot{\omega}_1 + (I_3 - I_2)\omega_2\omega_3 = M_1$$
  $$I_2\dot{\omega}_2 + (I_1 - I_3)\omega_3\omega_1 = M_2$$
  $$I_3\dot{\omega}_3 + (I_2 - I_1)\omega_1\omega_2 = M_3$$

  where M_1, M_2 and M_3 are the torque components in the body-fixed frame.

- I stress again that the Euler equations hold *only* if the body-fixed frame aligns with the principle axes. The most general equation of motion for rigid body rotation is the vector equation

  $$\boldsymbol{M} = \dot{\boldsymbol{L}} = \frac{\mathrm{d}[\mathbf{I}\boldsymbol{\omega}]}{\mathrm{d}t} = \mathbf{I}\dot{\boldsymbol{\omega}} + \boldsymbol{\omega} \times (\mathbf{I}\boldsymbol{\omega})$$

  where $\mathbf{I}$ is the body's inertia tensor about the point of rotation, $\boldsymbol{M}$ is the externally applied torque and $\boldsymbol{\omega}$ is the body's angular velocity. In the body-fixed system of principle axes, this equation simplifies to the Euler equations.

4.3 Motion of the Free Top

Discuss and analytically solve the equations of motion for a symmetric free top, then discuss the stability an asymmetric free top using a perturbation approach.

4.3.1 Symmetric Free Top

- A free top is fancy name for a freely rotating rigid body with no external torques acting on it. A *symmetric top* is a rotating rigid body with two equal principle moments are equal to each other e.g. $I_1 = I_2$, while the third moment I_3 is different from I_1 and I_2. The e_3 axis is the top's axis of symmetry, and is sometimes called the *figure axis.*

- For a free symmetric top with $I_1 = I_2$, the Euler equations become

$$I_1\dot{\omega}_1 = \omega_2\omega_3(I_1 - I_3)$$
$$I_1\dot{\omega}_2 = -\omega_3\omega_1(I_1 - I_3)$$
$$I_3\dot{\omega}_3 = 0$$

  ω_3 is the component (measured in the body-fixed system of principle axis) of the angular velocity about the axis of symmetry e_3 and, from the third Euler equation $I_3\dot{\omega}_3 = 0$, is constant for a free symmetric top.

- The components ω_1 and ω_2 are given by the coupled first-order LDEs

$$\dot{\omega}_1 = \Omega\omega_2 \qquad \text{and} \qquad \dot{\omega}_2 = -\Omega\omega_1 \qquad \text{where } \Omega = \frac{(I_1 - I_3)}{I_1}\omega_3$$

  The equations are solved by

$$\omega_1 = \omega_0\sin\Omega t \qquad \text{and} \qquad \omega_2 = \omega_0\cos\Omega t$$

  This solution has a nice physical interpretation: in the body-fixed frame, the free top's angular velocity vector $\boldsymbol{\omega}$ precesses about the figure axis e_3 with angular frequency Ω. The direction of $\boldsymbol{\omega}$'s rotation depends on the sign of Ω and thus on whether $I_1 > I_3$ or $I_3 > I_1$. The precession of $\boldsymbol{\omega}$ about the figure axis e_3 in the body-fixed frame is sometimes called *wobble.*

  I stress that this precession refers to motion of the angular momentum about the top's figure axis, as seen in the body fixed frame. We'll soon see another motion, also called precession, referring to motion of the top's figure axis around the lab's z axis $\tilde{e}_3$ as seen in the lab frame. Remember the two are different!

- What would we see in the lab frame? We know the angular momentum $\boldsymbol{L}$ is constant since the top is free. More so, ω_3 is constant, so $L_3 = I_3\omega_3$ is also constant. Since both $\boldsymbol{L}$ and L_3 are constant, the angle between the figure axis e_3 and angular momentum $\boldsymbol{L}$ is constant. The figure axis e_3 thus precesses about the fixed angular momentum $\boldsymbol{L}$, while the body rotates so that $\boldsymbol{\omega}$ remains between e_3 and $\boldsymbol{L}$. Click here for a nice animation.

  Again, this is *not* precession of the top's figure axis around the lab's z axis $\tilde{e}_3$ as seen in the lab frame. We'll see that shortly.

4.3.2   Free Symmetric Top In Terms of Euler Angles

Earlier we analyzed the free symmetric top in the *body-fixed frame* $\{e_a\}$. This is take two, using the Euler angles and working in the *lab* frame $\{\tilde{e}_a\}$.

- We choose the angular momentum $\mathbf{L}$ to align with the $\tilde{e}_3$ axis. The top is free, so $\boldsymbol{L}$ is conserved, as is the body-fixed frame component $L_3 = I_3\omega_3$; since both quantities are conserved, the angle between them is also conserved.

  The angle between L_3 and $\tilde{e}_3$ is precisely the second Euler angle, i.e. the nutation angle θ. Since this angle is constant, $\dot{\theta} = 0$ and the top does not perform nutation. In general, a free symmetric top does not perform nutation when $\boldsymbol{L}$ coincides with $\tilde{e}_3$.

- The next step is to use the Euler-angle expression for angular velocity in the body frame to solve for $\dot{\psi}$. Accounting for $\dot{\theta} = 0$, we have

$$\boldsymbol{\omega} = \omega_a e_a = \dot{\phi}\sin\theta\sin\psi e_1 + \dot{\phi}\sin\theta\cos\psi e_2 + (\dot{\phi}\cos\theta + \dot{\psi})e_3$$

  Equating the third components gives

$$\omega_3 = \dot{\phi}\cos\theta + \dot{\psi} \implies \dot{\psi} = \omega_3 - \dot{\phi}\cos\theta$$

- Next, we turn to the components ω_1 and ω_2. The previous section showed

$$\omega_1 = \omega_0\sin\Omega t \qquad \text{and} \qquad \omega_2 = \omega_0\cos\Omega t$$

  while the Euler angles tell us

$$\omega_1 = \dot{\phi}\sin\theta\sin\psi \qquad \text{and} \qquad \omega_2 = \dot{\phi}\sin\theta\cos\psi e_2$$

  Solving the system of equations leads to

$$\tan\Omega t = \tan\psi \implies \Omega t = \psi \implies \dot{\psi} = \Omega = \frac{(I_1 - I_3)}{I_1}\omega_3$$

- Combining the results $\dot{\psi} = \Omega$ and $\dot{\psi} = \omega_3 - \dot{\phi}\cos\theta$ leads to

$$\dot{\phi} = \frac{\omega_3 - \Omega}{\cos\theta} = \frac{I_3\omega_3}{I_1\cos\theta}$$

  which gives us the top's frequency of precession: $\dot{\phi} = \frac{I_3\omega_3}{I_1\cos\theta}$. This is the precession of the top's figure axis about the lab's z axis $\tilde{e}_3$ as seen in the lab frame, corresponding to the Euler angle ϕ.

4.3.3   Stability and Perturbations of The Asymmetric Top

- An asymmetric topic basically means a rigid body without any symmetries. This means none of the principal moments of inertia are equal, i.e. $I_1 \neq I_2 \neq I_3 \neq I_1$. The rotational motion of an asymmetric top is much more complicated than a free top's.

- We will analyze only a simple case: when the a free asymmetric top rotating about a principle axis of inertia, e.g. e_1 so $\boldsymbol{\omega}$ and e_1 have the same direction. In this case, in the body-fixed system of principle axes

$$w_1 = \Omega \qquad \text{and} \qquad \omega_2 = \omega_3 = 0$$

  This solves Euler's equations.

- We then ask what happens if the top's angular velocity is displaced slightly from the principle axis e_1, considering small perturbations about the angular velocity of the form

$$\omega_1 = \Omega + \eta_1 \qquad \omega_2 = \eta_2 \qquad \omega_3 = \eta_3$$

where the $\eta_a = \eta_a(t)$ and $\eta_a \ll \Omega$. We then substitute the ω_a into the Euler equations. Since the η_a are small, we ignore terms of order η^2 and higher, e.g. $\eta_2\eta_3 \approx 0$ in the first equation. The result is

$$I_1\dot{\eta}_1 = 0$$
$$I_2\dot{\eta}_2 = \Omega\eta_3(I_1 - I_3)$$
$$I_3\dot{\eta}_3 = \Omega\eta_2(I_2 - I_1)$$

Differentiating the second equation and solving for $\dot{\eta}_3$, then substituting $\dot{\eta}_3$ into the third equation gives

$$I_2\ddot{\eta}_2 = \frac{\Omega^2}{I_3}(I_1 - I_2)(I_3 - I_1)\eta_2 \equiv A\eta_2$$

Which has the familiar form $\ddot{x} = kx$. The stability of the perturbation depends on the sign of A:

- If $A < 0$ the disturbance will oscillate around the position of constant motion.
- If $A > 0$ the disturbance will grow exponentially, and the system is unstable.

From the definition $A \equiv \frac{\Omega^2}{I_3}(I_1 - I_2)(I_3 - I_1)$, the system is unstable ($A > 0$) if

$$I_2 < I_1 < I_3 \qquad \text{or} \qquad I_3 < I_1 < I_2$$

while all other possibilities are stable. Physically, this means a body will rotate stably about the principle axis with the largest or the smallest moment of inertia, but not about the intermediate axis.

4.4   The Heavy Symmetric Top

Analyze the basic dynamics of a heavy symmetric top with Lagrangian mechanics. Be sure to discuss the role of conserved quantities in simplifying the solution process. Discuss the conditions for uniform precession and a sleeping top.

A heavy top is a top acted on by gravity. Symmetric means that two of the principle moments of inertia are equal, so in the body-fixed system of principle axes we have $\mathbf{I} = (I_1, I_1, I_3)$. This analysis assumes the top is fixed at some point (e.g. on the ground) a distance l from the center of mass.

4.4.1   Lagrangian and Conserved Quantities

- The kinetic energy is $T = \frac{1}{2}I_1(\omega_1 + \omega_2)^2 + \frac{1}{2}I_3\omega_3^2$ and the potential energy is $V = mgl\cos\theta$. In terms of the Euler angles, the heavy symmetric top's Lagrangian is

$$\begin{aligned} L &= \frac{1}{2}I_1\left(\omega_1^2 + \omega_2^2\right) + \frac{1}{2}I_3\omega_3^2 - mgl\cos\theta \\ &= \frac{1}{2}I_1\left(\dot{\phi}^2\sin^2\theta + \dot{\theta}^2\right) + \frac{1}{2}I_3\left(\dot{\phi}\cos\theta + \dot{\psi}\right)^2 - mgl\cos\theta \end{aligned}$$

- Both ϕ and ψ are cyclic coordinates and lead to the conserved quantities

$$p_\phi = I_1 \sin^2\theta \, \dot\phi + I_3 \cos\theta(\dot\psi + \dot\phi\cos\theta)$$

$$p_\psi = I_3\left(\dot\psi + \cos\theta\,\dot\phi\right) = I_3\omega_3$$

  p_ψ is the top's angular momentum about the figure axis e_3. Like for the free symmetric top, the angular velocity component ω_3 about the figure axis is constant since no external torque acts along the figure axis.

Energy

- The top's total energy $E = T + V$ is also conserved, as for any isolated mechanical system. For future use, we define the constants

$$a = \frac{p_\psi}{I_1} = \frac{I_3}{I_1}\omega_3 \qquad b = \frac{p_\phi}{I_1}$$

  In terms of a and b, the derivatives $\dot\phi$ and $\dot\psi$ are

$$\dot\phi = \frac{b - a\cos\theta}{\sin^2\theta} \qquad \dot\psi = \frac{I_1 a}{I_3} - \frac{(b - a\cos\theta)\cos\theta}{\sin^2\theta}$$

  If $\theta = \theta(t)$ were known, we could theoretically integrate these two expressions to solve for $\phi(t)$ and $\psi(t)$. So the next step is to find $\theta(t)$.

Reduced Energy

- We define a reduced energy E' via $E = E' + \frac{1}{2}I_3\omega_3^2$. Since E, I_3 and ω_3 are constant, E' is also constant.

$$E' = \frac{1}{2}I_1\dot\theta^2 + V_{\text{eff}}(\theta), \qquad V_{\text{eff}}(\theta) = \frac{I_1(b - a\cos\theta)^2}{2\sin^2\theta} + mgl\cos\theta$$

  We then introduce a new coordinate $u = \cos\theta$, where $u \in [-1, 1]$ because $\theta \in [0, \pi]$, and two more constants

$$\alpha = \frac{2E'}{I_1} \qquad \text{and} \qquad \beta = \frac{2mgl}{I_1}$$

- The equations of motion for u, ϕ, ψ can now be written

$$\dot{u}^2 = (1 - u^2)(\alpha - \beta u) - (b - au)^2 \equiv f(u)$$

$$\dot\phi = \frac{b - au}{1 - u^2}$$

$$\dot\psi = \frac{I_1 a}{I_3} - \frac{u(b - au)}{1 - u^2}$$

  In theory, the equation for $\dot{u}^2$ could be solved analytically in terms of elliptic integrals, and then the time dependence of θ could be implicitly extracted using $u = \cos\theta$. With θ in hand, we could then solve for ϕ and ψ.

- But there's a better, more physical, approach: using the function $f(u)$ as a tool to analyze the nutation angle θ. Notice that $\dot{u}^2 = f(u)$ behaves as a cubic polynomial in u. A quick look at the limits $u \to \pm\infty$ shows

$$f(u) \to \begin{cases} +\infty \text{ as } u \to \infty \\ -\infty \text{ as } u \to -\infty \end{cases}$$

Also important is the behavior at ± 1:

$$f(\pm 1) = -(b \mp a)^2 \leq 0$$

Although $f(u)$ is mathematically defined on the entire real line, the region of *physical* interest is, happily, much smaller. Values of u are constrained to positive values of $f(u) = \dot{u}^2$ (corresponding to real values of $\dot{u}$) in the region $u \in [-1, 1]$ (because $|u| = |\cos\theta| \leq 1$).

We'll study three special cases of motion in the next three subsections.

4.4.2 Letting the Top Go: The Released Top

- In this case, sometime called the "released top", we spin the top about its axis of symmetry and let it go at some angle $\theta \neq 0$ with $\dot{\theta} = 0$. The initial conditions are

$$\dot{\theta}(0) = 0 \implies \dot{u}^2_{t=0} = 0 \implies f(u_{t=0}) = 0$$
$$\dot{\phi}(0) = 0 \implies b - a u_{t=0} = 0 \implies u_{t=0} = \frac{b}{a}$$

- Recall the quantity $p_\phi = I_1 \dot{\phi} \sin^2\theta + I_3 \omega_3 \cos\theta$ is a constant of motion. Evaluating p_ϕ at $t = 0$ when $\dot{\phi} = 0$ and $\theta = \theta_0$ gives the relationship

$$p_\phi = I_1 \dot{\phi} \sin^2\theta + I_3 \omega_3 \cos\theta = I_3 \omega_3 \cos\theta_0$$

This is enough to determine the qualitative motion of the top. When released, it starts to fall under the influence of gravity, so θ increases. As the top falls and θ increases (and thus $I_3 \omega_3 \cos\theta$ decreases) the term $I_1 \dot{\phi} \sin^2\theta$ must increase to keep p_ϕ constant, resulting as an increase in $\dot{\phi}$.

More so, to keep the signs of $\dot{\phi}$ and ω_3 balanced, the direction of precession must in the same as the direction of spin about the figure axis.

4.4.3 Uniform Precession Without Nutation: The Fast Top

- This means the tops spins uniformly without bobbing up and down. Mathematically, this means $\dot{\phi}$ is constant (uniform precession) and $\dot{\phi} = 0$ (no nutation).

Recall the relationship $\dot{\theta}^2 = f(u)$. To satisfy $\dot{\theta} = 0$ for all time, we need $f(u)$ to have only a single root u_0 and no other values of $f(u)$ in the physical range $u \in [-1, 1]$ and $u \geq 0$ which constrains the top's physical motion to a single point of constant $\dot{\theta} = 0$. The root must satisfy

$$f(u_0) = (1 - u_0^2)(\alpha - \beta u_0) - (b - a u_0)^2 = 0$$
$$f'(u_0) = -2u_0(\alpha - \beta u_0) - \beta(1 - u_0^2) + 2a(b - a u_0) = 0$$

Combining the equations gives the condition $\frac{1}{2}\beta = a\dot{\phi} - \dot{\phi}^2 u_0$, and substituting in the definitions $I_1 a = I_3 \omega_3$ and $\beta = \frac{2Mgl}{I_1}$ gives

$$Mgl = \dot{\phi}(I_3 \omega_3 - I_1 \dot{\phi} \cos\theta_0)$$

- Here is the physical interpretation: for a fixed value of ω_3 (the top's spin about its figure axis) and θ_0 (the angle from the vertical at which its released) there is an exact value of $\dot{\phi}$ (a "push" in the direction of precession) for which the top will spin without nutation, i.e. satisfying $Mgl = \dot{\phi}(I_3\omega_3 - I_1\dot{\phi}\cos\theta_0)$.

- Because the condition $Mgl = \dot{\phi}(I_3\omega_3 - I_1\dot{\phi}\cos\theta_0)$ is quadratic in $\dot{\phi}$, there are in principle two values of $\dot{\phi}$ for which the top could precess without nutation.

  If you rearrange and calculate the discriminant, you'll see equation has real solutions for $\dot{\phi}$ if

  $$\omega_3 > \frac{2}{I_3}\sqrt{MglI_1\cos\theta_0}$$

  In other words ω_3 must be large enough (the top must be spinning fast enough about its figure axis) to be able to precess without nutation at all. What happens when the top isn't spinning fast enough? It falls over!

4.4.4   The Sleeping Top

- In this case, we start the top spinning in an upright position with $\theta = \dot{\theta} = 0$. When the top spins completely upright with $\theta = 0$, it is called a *sleeping top*. But does the top stay there?

  Turning to the definitions $\dot{u}^2 = f(u)$ and $u = \cos\theta$, the sleeping top conditions $\theta = \dot{\theta} = 0$ are met if $u = 1$ (for $\theta = 0$) and $f(1) = 0$ (for $\dot{\theta} = 0$).

  In these conditions $a = b$ and $\alpha = \beta$. Here's a quick confirmation; feel free to skip.

  $$a\big|_{\theta=0} = \frac{I_3\omega_3\big|_{\theta=0}}{I_1} = \frac{I_3(\dot{\psi} + \dot{\phi})}{I_1} = \frac{p_\phi\big|_{\theta=0}}{I_1} = b\big|_{\theta=0}$$

  $$\alpha\big|_{\theta=0} = \frac{2E'\big|_{\theta=0}}{I_1} = \frac{2(0 + Mgl)}{I_1} = \beta$$

- With $a = b$ and $\alpha = \beta$ the function $f(u)$ simplifies to

  $$f(u) = (1 - u)^2\left[\alpha(1 - u) - a^2\right]$$

  so $f(u)$ has a double zero at $u = +1$, while a second zero occurs at $u_2 = 1 - \frac{a^2}{\alpha}$. There are two possibilities:

  - If $u_2 > 1$ (i.e. outside the region of physical motion) and $\omega_3^2 > \frac{4I_1Mgl}{I_3^2}$ the motion is stable: the only allowed physically allowed value of u corresponds to $\theta = 0$, since all other values of $f(u)$ are either negative or occur outside of $[-1, 1]$.
    For slight perturbations from the initial conditions, the nutation will stably oscillate with small displacements about $\theta = 0$.
  - If $u_2 < 1$ (inside the region of physical motion) and $\omega_3^2 < \frac{4I_1Mgl}{I_3^2}$, there is a region of allowed physical motion to the left of the zero $u = 1$. The motion is unstable: if the top is displaced slightly from $\theta = 0$, it will fall off into the region of decreasing u and thus increasing θ.

  Notice that the top is stable for fast spins ω_3 about the figure axis and unstable for slow spins. In practice, a sleeping top stays upright until ω_3 is slowed by friction to

  $$\omega_3 = \frac{4I_1Mgl}{I_3^2}$$

at which point the top starts to fall and precess.

5 Small Oscillations

5.1 Small Oscillations à la Tong

5.1.1 One Dimensional Example

- Let's start small, with a single particle and one degree of freedom q. In both Newtonian and Lagrangian mechanics, the motion is governed by a second-order differential equation with the general form

$$\ddot{q} = f(q)$$

for some (generally) non-linear function $f(q)$.

- At an equilibrium point $q = q_0$ the particle's acceleration $\ddot{q}$ is zero, so the equation of motion must satisfy $f(q_0) = 0$ at q_0.

  If we start at rest at equilibrium with the initial conditions

$$q = q_0 \qquad \text{and} \qquad \dot{q} = 0$$

the system will remain in equilibrium for all time (formally, solving the equation of motion gives $q(t) = q_0$).

- But what happens if we start ever so slightly away from q_0? Allowing for time-dependence, we write

$$q(t) = q_0 + \eta(t)$$

where $\eta(t)$ are small displacements about equilibrium q_0.

A first-order Taylor eqpansion of $\ddot{q} = f(q(t))$ gives

$$\ddot{q}(t) = f(q_0 + \eta(t)) \approx f(q_0) + f'(q_0)\eta(t) = f'(q_0)\eta(t)$$

where we've used the equilibrium requirement $f(q_0) = 0$. Since $\ddot{q} = \ddot{\eta}$, we have:

$$\ddot{\eta}(t) = f'(q_0)\eta(t)$$

- There are two possible outcomes for what happens next, depending on the sign of $f'(q_0)$.

  1. If $f'(q_0) < 0$, the acceleration $\ddot{\eta}$ points towards q_0, so q_0 is a position of stable equilibrium. The solution is sinusoidal oscillation about q_0:

$$\eta(t) = \eta_0 \cos(\omega(t - t_0)) \qquad \text{where} \quad \omega^2 = -f'(q_0)$$

  The system undergoes stable harmonic oscillations about q_0 at frequency ω.

  2. If $f'(q_0) > 0$, the acceleration $\ddot{\eta}$ points away from q_0, so q_0 is a position of unstable equilibrium. The solution is exponential:

$$\eta(t) = Ae^{\lambda t} + Be^{-\lambda t} \qquad \text{where} \quad \lambda = f'(q_0)$$

  for general initial conditions (besides the very special case $A = 0$) η grows rapidly and the first-order Taylor approximation breaks down. The system is said to be *linearly unstable* at q_0.

5.1.2 N Degrees of Freedom

- We now generalize to a system with N degrees of freedom. Our beginning steps are analogous to the one-dimensional case; the equations of motion are

$$\ddot{q}_i = f_i(q_1, \ldots q_N), \qquad i = 1, \ldots, n$$

where we've defined the vector of coordinates $\boldsymbol{q} = (q_1, \ldots, q_N)$. An equilibrium position $\boldsymbol{q}^0 = (q_1^0, \ldots q_N^0)$ satisfies

$$\ddot{q}_i = 0 \qquad \text{and} \qquad f_i(\boldsymbol{q}^0) \equiv f_i(q_1^0, \ldots, q_n^0) = 0 \qquad \text{for all } i = 1, \ldots, n$$

- Small displacements from the equilibrium point are written

$$q_i(t) = q_i^0 + \eta_i(t)$$

and a first-order Taylor expansion of $\ddot{\eta}_i = \ddot{q}_i$ gives

$$\begin{aligned}
\ddot{\eta}_i(t) &= f_i\left(q_1^0 + \eta_1(t), \ldots, q_N^0 + \eta_N(t)\right) \\
&\approx f_i\left(q_1^0, \ldots, q_N^0\right) + \sum_{j=1}^{N} \frac{\partial}{\partial q_j}\left[f_i\left(q_1^0, \ldots, q_N^0\right)\right]\eta_j \\
&= \sum_{j=1}^{N} \frac{\partial}{\partial q_j}\left[f_i\left(q_1^0, \ldots, q_N^0\right)\right]\eta_j \qquad \left(\text{since } f_i\left(q_1^0, \ldots, q_N^0\right) \equiv 0\right)
\end{aligned}$$

where the one-dimensional $f'(x_0)\eta$ is replaced with the total derivative $\frac{\partial f_i}{\partial q_j}\eta_j$ because $f_i = f_i(q_1, \ldots, q_N)$ is now a multi-variable function.

- Since were using multivariable calculus we'll switch to matrix form, which makes notation much cleaner. We introduce a vector of displacement $\boldsymbol{\eta} \in \mathbb{R}^N$ and the $(N \times N)$ matrix $\mathbf{F}$ given by

$$\boldsymbol{\eta} = \begin{bmatrix} \eta_1 \\ \vdots \\ \eta_n \end{bmatrix} \qquad \text{and} \qquad \mathbf{F} = \begin{bmatrix} \frac{\partial f_1}{\partial q_1} & \cdots & \frac{\partial f_1}{\partial q_N} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_N}{\partial q_1} & \cdots & \frac{\partial f_N}{\partial q_N} \end{bmatrix}_{q_i = q_i^0}$$

where $\mathbf{F}$ is the matrix of partial derivatives *evaluated at equilibrium* $q_i = q_i^0$. The system's equation of motion near equilibrium then reads

$$\ddot{\boldsymbol{\eta}} = \mathbf{F}\boldsymbol{\eta}$$

- We have reduced the dynamics of the system to an eigenvalue problem. To finish our solution we must $\mathbf{F}$'s eigenvectors and eigenvalues.

  But first, a necessary bit of linear algebra theory:

  - We cannot assume a priori that $\mathbf{F}$ has a complete set of orthogonal eigenvectors with real eigenvalues. It will turn out, however, that any matrix $\mathbf{F}$ derived from a physical Lagrangian system has all real eigenvalues. We'll prove this in a few paragraphs; for now we'll assume $\mathbf{F}$ has real eigenvalues.

– In general, any matrix, including e.g. $\mathbf{F}$, has different right and left eigenvectors $\boldsymbol{\mu}_a$ and $\boldsymbol{\zeta}_a$ satisfying

$$\mathbf{F}\boldsymbol{\mu}_a = \lambda_a^2 \boldsymbol{\mu}_a \qquad \boldsymbol{\zeta}_a^T \mathbf{F} = \lambda_a^2 \boldsymbol{\zeta}_a^T$$

where $a = 1, \ldots, N$ and *there is no implied sum over a.* The left and right eigenvectors are orthonormal and satisfy $\boldsymbol{\mu}_a \cdot \boldsymbol{\zeta}_b = \delta_{ab}$.

Note that even though the eigenvectors differ, the eigenvalues λ_a^2 are the same for $a = 1, \ldots, N$.

– Although the eigenvalues λ_a^2 are real for any physical system of interest, they are not always positive, allowing for the possibility of complex square roots $\sqrt{\lambda_a^2} = \pm\lambda_a \in \mathbb{C}$. That's okay; we'll see complex λ_a correspond to oscillation.

• Back to small oscillations, which we've reduced to the eigenvalue problem $\ddot{\boldsymbol{\eta}} = \mathbf{F}\boldsymbol{\eta}$. The most general solution to this equation is

$$\boldsymbol{\eta}(t) = \sum_a^N \boldsymbol{\mu}_a \left(A_a e^{\lambda_a t} + B_a e^{-\lambda_a t} \right)$$

where $\boldsymbol{\mu}_a$ are $\mathbf{F}$'s (left) eigenvectors, λ_a are $\mathbf{F}$'s eigenvalues, and A_a and B_a are integration constants determined by the initial conditions.

We have two possible behaviors for each a, depending on the sign on the eigenvalue λ_a^2

1. If $\lambda_a^2 < 0$ then $\pm\lambda_a \in \mathbb{C}$ is complex-valued and we write

$$\lambda_a = i\omega_a$$

for some real number $\omega_a \in \mathbb{R}$. This represents oscillation with the eigenfrequency ω_a about the equilibrium point q_a^N. The system is stable if displaced in the corresponding direction $\boldsymbol{\eta} = \boldsymbol{\mu}_a$.

2. If $\lambda_a^2 > 0$ then $\pm\lambda_a \in \mathbb{R}$ is real-valued (corresponding to an exponential solution $\eta_a(t) = A_a e^{\lambda_a t} + B_a e^{-\lambda_a t}$) and the system has a linear instability if displaced in the corresponding direction $\boldsymbol{\eta} = \boldsymbol{\mu}_a$.

• The eigenvectors $\boldsymbol{\mu}_a$ are called *normal modes* and the associated value ω_a is called the *eigenfrequency* or *normal frequency*. The equilibrium point $\boldsymbol{q}^0 = (q_1^0, \ldots, q_N^0)$ is stable only if $\lambda_a^2 < 0$ for all $i = 1, \ldots, n$, in which case the system oscillates about the equilibrium point as a linear superposition of the normal modes, each at the corresponding eigenfrequency.

In real form, the solution can be written

$$\boldsymbol{\eta}(t) = \sum_{a;\lambda_a^2>0} \boldsymbol{\mu}_a \left[A_a e^{\lambda_a t} + B_a e^{-\lambda_a t} \right] + \sum_{a;\lambda_a^2<0} A_a \boldsymbol{\mu}_a \cos(\omega_a(t - t_a))$$

where A_a, B_a and t_a are $2n$ integration constants determined by the initial conditions; the t_a replace the B_a for the oscillatory modes with $\lambda_a^2 < 0$.

Proof: F's Eigenvectors are Real

We want to show $\mathbf{F}$'s eigenvalues are real as long as $\mathbf{F}$ is derived from a physical Lagrangian system (i.e. something we'd actually see in nature).

- Consider a general Lagrangian of the form

$$L = \frac{1}{2} \sum_{i,j} \mathrm{T}_{ij}(\boldsymbol{q}) \dot{q}_i \dot{q}_j - V(\boldsymbol{q})$$

where $\mathbf{T} : \mathbb{R}^N \to \mathbb{R}^{N \times N}$ is matrix encoding the coefficients of the kinetic energy and $\boldsymbol{q} \in \mathbb{R}^N$ is a vector of coordinates. We'll assume $\mathbf{T}$ is invertible and positive definite for all $\boldsymbol{q}$. Finally, we'll need the potential energy matrix $\mathbf{V} : \mathbb{R}^N \to \mathbb{R}^{N \times N}$

$$\mathbf{V}(\boldsymbol{q}) = \begin{bmatrix} \frac{\partial^2 V}{\partial q_1 \partial q_1} & \cdots & \frac{\partial^2 V}{\partial q_1 \partial q_N} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 V}{\partial q_1 \partial q_N} & \cdots & \frac{\partial^2 V}{\partial q_N \partial q_N} \end{bmatrix}$$

Both $\mathbf{T}$ and $\mathbf{V}$ come out real-valued and symmetric by construction.

- Expanding about the equilibrium point $\boldsymbol{q}^0$ to linear order in displacement $\boldsymbol{\eta}$ gives

$$\mathbf{T}(\boldsymbol{q}^0)\ddot{\boldsymbol{\eta}} = -\mathbf{V}(\boldsymbol{q}^0)\boldsymbol{\eta} \qquad \text{or} \qquad \mathrm{T}_{ij}(\boldsymbol{q}^0)\ddot{\eta}_j = -\mathrm{V}_{ij}(\boldsymbol{q}^0)\eta_j \quad \text{for } i = 1, \dots, N$$

This might look frightening at first, but it is really just a set of equations of motion for the displacements $\ddot{\boldsymbol{\eta}}(t)$.

- We'll always be working with $\mathbf{T}$ and $\mathbf{V}$ evaluated at equilibrium $\boldsymbol{q}^0$, so let's define the shorthand

$$\mathbf{T}^0 = \mathbf{T}(\boldsymbol{q}^0) \qquad \text{and} \qquad \mathbf{V}^0 = \mathbf{V}(\boldsymbol{q}^0)$$

Rearranging our result to get $\ddot{\boldsymbol{\eta}} = -\left(\mathbf{T}^0\right)^{-1} \mathbf{V}^0 \boldsymbol{\eta}$ and comparing to the equation of motion $\ddot{\boldsymbol{\eta}} = \mathbf{F}\boldsymbol{\eta}$, we see

$$\mathbf{F} = -\left(\mathbf{T}^0\right)^{-1} \mathbf{V}^0$$

- While $\mathbf{T}^0$ and $\mathbf{V}^0$ are both symmetric, they are not necessarily simultaneously diagonalizable (this occurs only if $\mathbf{T}^0$ is a scalar multiple of the identity matrix $\mathbf{I}$), which means that $\mathbf{F}$ is not symmetric in general, as mentioned in the earlier interlude on linear algebra theory.

  However, we can show $\boldsymbol{F}$ has real eigenvalues. Since $\mathbf{F} = -\left(\mathbf{T}^0\right)^{-1} \mathbf{V}^0$ we have

$$\mathbf{F}\boldsymbol{u} = \lambda^2 \boldsymbol{\mu} \implies \mathbf{V}^0 \boldsymbol{\mu} = -\lambda^2 \mathbf{T}^0 \boldsymbol{\mu}$$

  Multiplying the equation by complex conjugate $\bar{\boldsymbol{\mu}}$ gives

$$\bar{\boldsymbol{\mu}} \mathbf{V}^0 \boldsymbol{\mu} = -\lambda^2 \bar{\boldsymbol{\mu}} \mathbf{T}^0 \boldsymbol{\mu}$$

  Since $\mathbf{V}^0$ and $\mathbf{T}^0$ are both symmetric, the quantities $\bar{\boldsymbol{\mu}} \mathbf{V}^0 \boldsymbol{\mu}$ and $\bar{\boldsymbol{\mu}} \mathbf{T}^0 \boldsymbol{\mu}$ are both real, which suggests λ^2 must be real too.

  In principle if both matrix quantities were zero the trivial equality $0 = 0$ would hold even for $\lambda^2 \in \mathbb{C}$. But this can't happen, since we've assumed $\mathbf{T}$ is positive definite we know $\bar{\boldsymbol{\mu}} \mathbf{T}^0 \boldsymbol{\mu} > 0$. It follows that $\mathbf{F}$'s eigenvalues λ^2 are indeed real.

5.2   Small Oscillations à la Goldstein

5.2.1   Direct Lagrangian Mechanics Approach

- We consider a system of N degrees of freedom with the coordinates $\boldsymbol{q} = (q_1, \ldots, q_N)$ whose Lagrangian can be written

$$L = T - V = \frac{1}{2} \sum_{i,j} \mathrm{T}_{ij}(\boldsymbol{q}) \dot{q}_i \dot{q}_j - V$$

where $\mathbf{T}$ is an $(N \times N)$ matrix encoding the coefficients of the kinetic energy. By construction $\mathbf{T}$ turns out to be symmetric, so $\mathrm{T}_{ij} = \mathrm{T}_{ji}$. We'll assume $\frac{\partial L}{\partial t} = 0$, so the total energy is conserved and can be written

$$H = \frac{\partial L}{\partial \dot{q}_i} \dot{q}_i - L = T + V$$

- We assume there exists a stationary solution of Lagrange's equations at the equilibrium position

$$\boldsymbol{q}(t) = \boldsymbol{q}^0 \qquad \text{or, by components} \qquad q_i(t) = q_i^0 \quad \text{for } i = 1, \ldots, N$$

At such a point the system is stationary so $\dot{q}_i = 0$ for all i and the kinetic energy term vanishes from the Lagrangian. We have

$$L\big|_{\boldsymbol{q}^0} = -V\big|_{\boldsymbol{q}^0} \qquad \text{and} \qquad \frac{\partial L}{\partial q_i}\bigg|_{\boldsymbol{q}^0} = -\frac{\partial V}{\partial q_i}\bigg|_{\boldsymbol{q}^0} = 0$$

- Stationary states $\boldsymbol{q}^0$ can be absolutely stable, unstable, or metastable.

  - *Absolutely stable states* satisfy

    $$V(\boldsymbol{q}^0) < \boldsymbol{V}(\boldsymbol{q}) \quad \text{for all } \boldsymbol{q} \neq \boldsymbol{q}^0$$

    In other words, absolutely stable states are positions of absolute minimum of the potential energy V.

  - *Metastable states* must satisfy the same condition, but only in a neighborhood of $\boldsymbol{q}^0$

    $$V(\boldsymbol{q}^0) < \boldsymbol{V}(\boldsymbol{q}) \quad \text{for } \boldsymbol{q} \text{ near } \boldsymbol{q}^0$$

    In other words, metastable states are local minima of the potential energy V.

  - *Unstable states* satisfy $V(\boldsymbol{q}^0) > \boldsymbol{V}(\boldsymbol{q})$. Such states are maxima of the potential energy, and the force $\boldsymbol{F} = -\nabla V$ pushes particles away from the equilibrium position. Such states are not of interest in the study of small oscillations.

- We'll deal with metastable and absolutely stable states. We denote small displacements from equilibrium by $\boldsymbol{\eta}(t)$ and write

$$\boldsymbol{q}(t) = \boldsymbol{q}^0 + \boldsymbol{\eta}(t) \qquad \text{or, by components} \qquad q_i(t) = q_i^0(t) + \eta_i(t)$$

We then assume $V(\boldsymbol{q})$ is an analytic function of the coordinates $\boldsymbol{q}$, which is just a guarantee we can expand it into a Taylor series about equilibrium $\boldsymbol{q}^0$. To second

order in $\boldsymbol{q}$ we have

$$V(\boldsymbol{q}) \approx V(\boldsymbol{q}^0) + \left.\frac{\partial V}{\partial q_i}\right|_{\boldsymbol{q}^0} \eta_i + \frac{1}{2}\left.\frac{\partial^2 V}{\partial q_i \partial q_j}\right|_{\boldsymbol{q}^0} \eta_i \eta_j$$

$$= V(\boldsymbol{q}^0) + \frac{1}{2}\left.\frac{\partial^2 V}{\partial q_i \partial q_j}\right|_{\boldsymbol{q}^0} \eta_i \eta_j$$

where the linear term vanishes by the equilibrium condition $\left.\frac{\partial V}{\partial q_i}\right|_{\boldsymbol{q}^0} = 0$. The metastable equilibrium condition $V(\boldsymbol{q}) > V(\boldsymbol{q}^0)$ then implies

$$\frac{1}{2}\left.\frac{\partial^2 V}{\partial q_i \partial q_j}\right|_{\boldsymbol{q}^0} \eta_i \eta_j > 0 \iff \frac{1}{2}V_{ij}\eta_i\eta_j > 0$$

where the second expression is written in terms of the matrix $\mathbf{V}$ of potential energy second derivatives

$$\mathbf{V} = \begin{bmatrix} \frac{\partial^2 V}{\partial q_1 \partial q_1} & \cdots & \frac{\partial^2 V}{\partial q_1 \partial q_N} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 V}{\partial q_1 \partial q_N} & \cdots & \frac{\partial^2 V}{\partial q_N \partial q_N} \end{bmatrix}_{\boldsymbol{q}=\boldsymbol{q}^0}$$

By construction $\mathbf{V}$ comes out to be symmetric and positive-definite, which will be important later on.

- We can then write the system's Lagrangian near equilibrium in the matrix notation

$$L = \frac{1}{2}T_{ij}\dot{\eta}_i\dot{\eta}_j - \frac{1}{2}V_{ij}\eta_i\eta_j + V^0$$

where $\mathbf{T} = \left.\mathbf{T}\right|_{\boldsymbol{q}^0}$. Lagrange's equations for the displacements η_i are then written

$$T_{ij}\ddot{\eta}_j + V_{ij}\eta_j = 0 \qquad \text{or} \qquad \mathbf{T}\ddot{\boldsymbol{\eta}} + \mathbf{V}\boldsymbol{\eta} = 0$$

- The basic solution of the equations of motion $T_{ij}\ddot{\eta}_j + V_{ij}\eta_j = 0$ are so called *characteristic oscillations*, which are harmonic oscillations of the all the displacement components $\eta_i(t)$ with the same frequency ω, represented by the ansatz

$$\eta_j(t) = \alpha a_j e^{-i\omega t}$$

where α describes the amplitude of the oscillations and the unit vector $\boldsymbol{a} = (a_1, \ldots, a_N)$ encodes the direction of each displacement component η_i.

To get a relationship between ω^2 and $\boldsymbol{a}$, we calculate $\ddot{\eta}_j = -\omega^2 \alpha a_j e^{-\omega t}$, plug this in to the equation of motion, cancel like terms and rearrange to get

$$V_{ij}a_j = \omega^2 T_{ij}a_j \qquad \text{or} \qquad \mathbf{V}\boldsymbol{a} = \omega^2 \mathbf{T}\boldsymbol{a}$$

- The result $\mathbf{V}\boldsymbol{a} = \omega^2 \mathbf{T}\boldsymbol{a}$ is important: we have reduced the dynamics of small oscillations to a generalized eigenvalue problem involving the matrices $\mathbf{T}$ and $\mathbf{V}$. Because the matrices are symmetric, the eigenvalues ω^2 are all real and because $\mathbf{T}$ is positive definite, we know the eigenvalues will be positive $\omega^2 > 0$.

Note, however, that this is in general *not* a simple eigenvalue problem of the form $\mathbf{A}\boldsymbol{x} = \omega^2\boldsymbol{x}$; there are two matrices involved!

5.2.2   The Simplified Eigenvalue Problem

- Let's start on more familiar territory with the simple eigenvalue problem of the general form $\mathbf{A}\boldsymbol{x} = \omega^2 \boldsymbol{x}$. This occurs if $\mathbf{T} = T\mathbf{I}$, i.e. if $\mathbf{T}$ is diagonal with identical scalar elements T. In this case, the problem reduces to the familiar eigenvalue problem $\mathbf{V}\boldsymbol{a} = \omega^2 \mathbf{T}\boldsymbol{a}$.

- Because $\boldsymbol{V}$ is symmetric (and the $\mathbf{a}$ are normalized by construction) the eigenvectors are orthogonal:

$$\boldsymbol{a}_i^T \cdot \boldsymbol{a}_j = \delta_{ij}$$

We write the eigenvectors $\boldsymbol{a}_i$ in the $N \times N$ matrix $\mathbf{A}$

$$\mathbf{A} = [\boldsymbol{a}_1, \ldots, \boldsymbol{a}_N] = \begin{bmatrix} a_{11} & \cdots & a_{1N} \\ \vdots & \ddots & \vdots \\ a_{N1} & \cdots & a_{NN} \end{bmatrix}$$

Because the $\boldsymbol{a}_i$ are orthonormal (i.e. $\boldsymbol{a}_i \cdot \boldsymbol{a}_j = \delta_{ij}$) we have

$$\mathbf{A}^T \mathbf{A} = \mathbf{I} \implies \mathbf{A} = \mathbf{A}^{-1}$$

- Some new notation coming fast: for $\mathbf{V}$'s eigenvalues we'll write $\omega^2 \equiv \lambda$, so the eigenvalue problem reads

$$\mathbf{V}\boldsymbol{a} = \lambda \mathbf{T}\boldsymbol{a} = \lambda T\mathbf{I} \qquad \text{or} \qquad (\mathbf{V} - \lambda T\mathbf{I})\boldsymbol{a} = 0$$

(We're still assuming $\mathbf{T} = T\mathbf{I}$). $\mathbf{A}$'s eigenvalues are $\lambda_i T$ for $i = 1, \ldots, N$, which we'll write $\tilde{\lambda}_i \equiv \lambda_i T$ for shorthand, so our problem now reads $(\mathbf{V} - \tilde{\lambda}\mathbf{I})\boldsymbol{a} = 0$. Since we require $\boldsymbol{a} \neq \mathbf{0}$, the equation is satisfied if $|\mathbf{V} - \lambda T\mathbf{I}| = 0$.

- In terms of left and right eigenvectors, the eigenvalue problem is

$$(\mathbf{V} - \tilde{\lambda}_i\mathbf{I})\boldsymbol{a}_i = 0 \qquad \text{and} \qquad \boldsymbol{a}_j^T(\mathbf{V} - \tilde{\lambda}_j\mathbf{I}) = 0$$

We then combine these two equations into the matrix equation

$$\mathbf{A}^T \mathbf{V}\mathbf{A} - \mathbf{A}^T \mathbf{A}\boldsymbol{\Lambda} = 0$$

where $\boldsymbol{\Lambda}$ is a diagonal matrix whose diagonal elements are the eigenvalues $\tilde{\lambda}_i$. Because the eigenvectors $\boldsymbol{a}_i$ are orthonormal and $\mathbf{A}^T \mathbf{A} = \mathbf{I}$, the matrix $\boldsymbol{\Lambda}$ is also unitary.

From the matrix equation $\mathbf{A}^T \mathbf{V}\mathbf{A} = \mathbf{A}^T \mathbf{A}\boldsymbol{\Lambda}$, we can interpret the eigenvector matrix $\mathbf{A}$ as a transformation matrix that transforms $\mathbf{V}$ into the diagonal form $\boldsymbol{\Lambda}$, i.e.

$$\mathbf{A}^T \mathbf{V}\mathbf{A} = \mathbf{A}^{-1}\mathbf{V}\mathbf{A} = \mathbf{A}^{-1}\mathbf{A}\boldsymbol{\Lambda} = \boldsymbol{\Lambda}$$

where we've used $\mathbf{V}\mathbf{A} = \mathbf{A}\boldsymbol{\Lambda}$ and the orthogonality property $\mathbf{A}^T = \mathbf{A}^{-1}$.

5.2.3   Generalized Eigenvalue Problem

- On to the generalized eigenvalue problem with $\mathbf{T} \neq T\mathbf{I}$

$$\mathbf{V}\boldsymbol{a} = \omega^2 \mathbf{T}\boldsymbol{a} = \lambda \mathbf{T}\boldsymbol{a}$$

where we'll continue with the notation $\lambda \equiv \omega^2$ for the eigenvalues. Both $\mathbf{V}$ and $\mathbf{T}$ are real and symmetric, so the eigenvalues λ and eigenvectors $\boldsymbol{a}$ are real.

- As before, we write $(\mathbf{V} - \lambda\mathbf{T})\boldsymbol{a} = 0$ and require $|\mathbf{V} - \lambda\mathbf{T}| = 0$, which disregards the trivial solution $\boldsymbol{a} = 0$.

  In terms of left and right eigenvectors, the eigenvalue problem reads

  $$\mathbf{V}\boldsymbol{a}_i = \lambda_i \mathbf{T}\boldsymbol{a}_i \qquad \text{and} \qquad \mathbf{a}_j^T \mathbf{V} = \lambda_j \mathbf{a}_j^T \mathbf{T}$$

  Next, we'll prove the eigenvectors satisfy a generalized orthogonality. Multiplying the first equation from the left by $\boldsymbol{a}_j^T$ and the second equation from the right by $\boldsymbol{a}_i$ gives

  $$\mathbf{a}_j^T \mathbf{V}\boldsymbol{a}_i = \lambda_i \mathbf{a}_j^T \mathbf{T}\boldsymbol{a}_i \qquad \text{and} \qquad \mathbf{a}_j^T \mathbf{V}\boldsymbol{a}_i = \lambda_j \mathbf{a}_j^T \mathbf{T}\boldsymbol{a}_i$$

  Subtracting the equations leads to $(\lambda_i - \lambda_j)\mathbf{a}_j^T \mathbf{T}\boldsymbol{a}_i = 0$, which implies

  $$\mathbf{a}_j^T \mathbf{T}\boldsymbol{a}_i = 0 \qquad \text{for} \quad \lambda_i \neq \lambda_j$$

  Even for degenerate eigenvalues $\lambda_i = \lambda_j, i \neq j$, we can choose the eigenvectors so they satisfy the orthogonality relation. More so, because $\mathbf{T}$ is positive definite, its diagonal elements must be positive and nonzero. This allows us to chose the eigenvectors to satisfy the generalized normalization condition

  $$\boldsymbol{a}_i^T \mathbf{T}\boldsymbol{a}_i = 1$$

- Together, the generalized orthogonality ($\mathbf{a}_j^T \mathbf{T}\boldsymbol{a}_i = 0$) and normalization ($\boldsymbol{a}_i^T \mathbf{T}\boldsymbol{a}_i = 1$) conditions mean that the eigenvector matrix $\mathbf{A}$ transforms the kinetic energy matrix $\mathbf{T}$ into the identity matrix

  $$\mathbf{A}^T \mathbf{T}\mathbf{A} = \mathbf{I}$$

  Like for the simpler case, we can now write the eigenvalue problem as a matrix equation

  $$\mathbf{V}\mathbf{A} = \mathbf{T}\mathbf{A}\boldsymbol{\Lambda}$$

  where $\boldsymbol{\Lambda}$ is a diagonal matrix whose elements are the eigenvalues λ_i.

  As before, we can view $\mathbf{A}$ as a transformation matrix that transforms $\mathbf{V}$ into the diagonal form $\boldsymbol{\Lambda}$. Using $\mathbf{V}\mathbf{A} = \mathbf{T}\mathbf{A}\boldsymbol{\Lambda}$ and $\mathbf{A}^T \mathbf{T}\mathbf{A} = \mathbf{I}$ we have

  $$\mathbf{A}^T \mathbf{V}\mathbf{A} = \mathbf{A}^T \mathbf{T}\mathbf{A}\boldsymbol{\Lambda} = \boldsymbol{\Lambda}$$

  In the general case, however, $\mathbf{A}$ has the additional property of transforming $\mathbf{T}$ into the identity matrix $\mathbf{I}$.

5.2.4   Normal Coordinates

- A few sections ago we derived the equations of motion governing small oscillations

  $$\mathrm{T}_{ij}\ddot{\eta}_j + \mathrm{V}_{ij}\eta_j = 0$$

  and said the basic solutions were the characteristic oscillations

  $$\eta_j(t) = \alpha a_j e^{-i\omega t}$$

  The general solution is a linear superposition of the characteristic oscillations.

- The general solution is best written in terms of *normal coordinates*, which replace the small displacements η_i with the amplitudes α_j of the characteristic oscillations $\boldsymbol{a}_j$. We write the transformation to the normal coordinates as

$$\eta_i(t) = \alpha_j(t) a_{j_i} = a_{j_i} \alpha_j(t)$$

where a_{j_i} is the i-th component of the j-th eigenvector $\boldsymbol{a}_j$. Note also that a_{j_i} corresponds to the element A_{ij} of the eigenvector matrix $\mathbf{A} = [\boldsymbol{a}_1, \ldots, \boldsymbol{a}_N]$.

  We can thus write the transformation to normal coordinates in the matrix form

$$\boldsymbol{\eta}(t) = \mathbf{A}\boldsymbol{\alpha}(t) \qquad \text{or} \qquad \boldsymbol{\eta}^T(t) = \boldsymbol{\alpha}^T(t)\mathbf{A}^T$$

- The normal coordinates allow us to concisely write the system's Lagrangian

$$L = \frac{1}{2}\left(T_{ij}\dot{\eta}_i\dot{\eta}_j - V_{ij}\eta_i\eta_j\right) + V^0 = \frac{1}{2}\left(\dot{\boldsymbol{\eta}}^T\mathbf{T}\dot{\boldsymbol{\eta}} - \boldsymbol{\eta}^T\mathbf{V}\boldsymbol{\eta}\right) + V^0$$

in terms of the eigenvalues ω_j^2. First, let's redefine the potential so $V^0 = 0$. Then, using $\boldsymbol{\eta} = \mathbf{A}\boldsymbol{\alpha}$ and the identities $\mathbf{A}^T\mathbf{T}\mathbf{A} = \mathbf{I}$ and $\mathbf{A}^T\mathbf{V}\mathbf{A} = \boldsymbol{\Lambda} = \operatorname{diag}(\omega_j^2)$ we have

$$\begin{aligned} L &= \frac{1}{2}\left(\dot{\boldsymbol{\eta}}^T\mathbf{T}\dot{\boldsymbol{\eta}} - \boldsymbol{\eta}^T\mathbf{V}\boldsymbol{\eta}\right) = \frac{1}{2}\left(\dot{\boldsymbol{\alpha}}^T\mathbf{A}^T\mathbf{T}\mathbf{A}\dot{\boldsymbol{\alpha}} - \boldsymbol{\alpha}^T\mathbf{A}^T\mathbf{V}\mathbf{A}\boldsymbol{\alpha}\right) \\ &= \frac{1}{2}\left(\dot{\boldsymbol{\alpha}}^T\mathbf{I}\dot{\boldsymbol{\alpha}} - \boldsymbol{\alpha}^T\boldsymbol{\Lambda}\boldsymbol{\alpha}\right) = \frac{1}{2}\left(\dot{\boldsymbol{\alpha}}^T\dot{\boldsymbol{\alpha}} - \boldsymbol{\alpha}^T\boldsymbol{\Lambda}\boldsymbol{\alpha}\right) \\ &= \frac{1}{2}\sum_j\left(\dot{\alpha}_j^2 - \omega_j^2\alpha_j^2\right) \end{aligned}$$

- The time dependence of each normal coordinate α_j is the harmonic oscillation

$$\alpha_j(t) = \alpha_{j_0}\cos(\omega_j t - \delta)$$

where the amplitude α_{j_0} and phase δ are integration constants determined by the initial conditions.

  The Lagrangian is thus a sum of $j = 1, \ldots, N$ uncoupled harmonic oscillators $\alpha_j(t)$, each vibrating at the frequency ω_j, where the amplitude of oscillation is the j-th normal coordinate α_j.

6 Hamiltonian Mechanics

6.1 The Hamiltonian Formalism

Give an overview of the Hamiltonian formulation of mechanics. Be sure to discuss the relationship between the Lagrangian and Hamiltonian formalism via the Legendre transform. Discuss conserved quantities and the least action principle in the Language of Hamiltonian mechanics.

6.1.1 Basics of the Hamiltonian Formalism

- The foundation of the Lagrangian formalism is the Lagrangian function $L(q_i, \dot{q}_i, t)$ where $q_i, i = 1, \ldots, N$ are the N generalized coordinates describing the system. The equations of motion are

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{q}_i}\right) - \frac{\partial L}{\partial q_i} = 0$$

  In the Lagrangian formalism the equations of motion are N second order differential equations; for N second order equations we need $2N$ initial conditions to define our system; typically there are $q_i(0)$ and $\dot{q}_i(0)$.

- The Hamiltonian formalism also uses N generalized coordinates q_i. However, instead of the derivatives $\dot{q}_i$, Hamiltonian mechanics uses the *generalized momenta*

$$p_i \equiv \frac{\partial L}{\partial \dot{q}_i}$$

  Because they are derived from the Lagrangian, the generalized momenta are also functions of $q_i, \dot{q}_i$ and t; we have $p_i = p_i(q_j, \dot{q}_j, t)$.

- In terms of the generalized momenta $p_i = \frac{\partial L}{\partial \dot{q}_i}$ the Lagrangian equations of motion read

$$\frac{\mathrm{d}p_i}{\mathrm{d}t} - \frac{\partial L}{\partial q_i} = 0 \implies \dot{p}_i = \frac{\partial L}{\partial q_i}$$

  Notice we have eliminated the derivatives $\dot{q}_i$ in favor of the momenta p_i.

- Recall the set $\{q_i\}$ defines a point in our system's n-dimensional configuration space C, and the time evolution of our system is represented by an n-dimensional path in C.

  However, the *state* of the system is defined by both $\{q_i\}$ and $\{p_i\}$ in the sense that knowing both the coordinates and momenta allows us to determine the system's state at all future times. The pair $\{q_i, p_i\}$ represents a point in the system's $2N$-dimensional *phase space*. Knowing a single point in phase space uniquely determines the system's evolution in time. Since a point *uniquely* defines the system's future, paths in phase space never cross.

6.1.2 Mathematical Detour: The Legendre Transform

We want to find a function defined on a system's phase space that will determine the unique evolution of q_i and p_i. Naturally, such a function must be a function of q_i and p_i, but must also contain the same information as the Lagrangian $L(q_i, \dot{q}_i, t)$ (which we know also determines the system's time evolution). We create our desired function with a mathematical technique called the *Legendre transform*. The Legendre transform goes something as follows:

- First, consider an arbitrary function $f(x, y)$, so the total derivative is

$$\mathrm{d}f = \frac{\partial f}{\partial x}\,\mathrm{d}x + \frac{\partial f}{\partial y}\,\mathrm{d}y$$

  Next we define a new function $g(x, y, u) = ux - f(x, y)$ of three variables x, y and u. The total derivative $\mathrm{d}g$ is

$$\mathrm{d}g = \mathrm{d}[ux] - \mathrm{d}f = u\,\mathrm{d}x + x\,\mathrm{d}u - \frac{\partial f}{\partial x}\,\mathrm{d}x - \frac{\partial f}{\partial y}\,\mathrm{d}y$$

- The total derivative simplifies considerably for a special choice of u. It is $u(x, y) = \frac{\partial f}{\partial x}$. In this case the terms $u\,\mathrm{d}x$ and $\frac{\partial f}{\partial x}\,\mathrm{d}x$ vanish. We have

$$\mathrm{d}g = x\,\mathrm{d}u - \frac{\partial f}{\partial y}\,\mathrm{d}y$$

  Notice that all instances of $\mathrm{d}x$ are gone; only $\mathrm{d}u$ and $\mathrm{d}y$ remain. In other words, g is naturally thought of as a function of only u and y, i.e. $g = g(u, y)$.

- To get an explicit expression for $g(u, y)$, we first solve for x by inverting $u(x, y) = \frac{\partial f}{\partial x}$ to get $x = x(u, y)$, then insert this into the definition $g = ux - f$ to get

$$g(u, y) = ux(u, y) - f(x(u, y), y)$$

  This last result is the *Legendre transform*. It takes us from the function $f(x, y)$ to the function $g(u, y)$, eliminating the dependence on x.

  The Legendre transform g preserves all information about the original f. Concretely, we can recover $f(x, y)$ from $g(u, y)$ and $u(x, y) = \frac{\partial f}{\partial x}$ with

$$\left.\frac{\partial g}{\partial u}\right|_y = x(u, y) \qquad \text{and} \qquad \left.\frac{\partial g}{\partial y}\right|_u = -\frac{\partial f}{\partial y}$$

  which means the inverse Legendre transform

$$f(x, y) = \left(\frac{\partial g}{\partial u}\right)u - g(u, y)$$

  recovers the original function f.

6.1.3   Hamilton's Equations of Motion

- In Hamiltonian mechanics, the Lagrangian $L(q, q_i, t)$ is replaced by the *Hamiltonian*, which is a Legendre transform of L to eliminate $\dot{q}_i$. H is defined as

$$H(q_i, p_i, t) = p_i \dot{q}_i - L(q_i, \dot{q}_i, t)$$

  To explicitly see the elimination of $\dot{q}_i$, we calculate the total derivative $\mathrm{d}H$:

$$\mathrm{d}H = p_i\,\mathrm{d}\dot{q}_i + \dot{q}_i\,\mathrm{d}p_i - \left(\frac{\partial L}{\partial q_i}\,\mathrm{d}q_i + \frac{\partial L}{\partial \dot{q}_i}\,\mathrm{d}\dot{q}_i + \frac{\partial L}{\partial t}\,\mathrm{d}t\right)$$

  From the construction $p_i = \frac{\partial L}{\partial \dot{q}_i}$, the terms $p_i\,\mathrm{d}\dot{q}_i$ and $-\frac{\partial L}{\partial \dot{q}_i}\,\mathrm{d}\dot{q}_i$ cancel. We have

$$\mathrm{d}H = \dot{q}_i\,\mathrm{d}p_i - \frac{\partial L}{\partial q_i}\,\mathrm{d}q_i - \frac{\partial L}{\partial t}\,\mathrm{d}t$$

  and indeed we see $H = H(q_i, p_i, t)$.

- More so, we can compare our expression for $\mathrm{d}H$ to the general total differential $\mathrm{d}H = \frac{\partial H}{\partial q_i}\,\mathrm{d}q_i + \frac{\partial H}{\partial p_i}\,\mathrm{d}p_i + \frac{\partial H}{\partial t}\,\mathrm{d}t$ to conclude

$$\frac{\partial H}{\partial p_i} = \dot{q}_i \qquad \frac{\partial H}{\partial t} = -\frac{\partial L}{\partial t} \qquad \frac{\partial H}{\partial q_i} = -\frac{\partial L}{\partial q_i} = -\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{q}_i}\right) \equiv -\dot{p}_i$$

where we use the Euler-Lagrange equations $\frac{\partial L}{\partial q_i} = \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{q}_i}\right)$ in the last equality.

These last results are important: they are the equations of motion in the Hamiltonian formalism. They are most commonly written

$$\dot{q}_i = \frac{\partial H}{\partial p_i} \qquad \text{and} \qquad \dot{p}_i = -\frac{\partial H}{\partial q_i}$$

Note that the Hamiltonian formalism replaces the Lagrangian's n second order differential equations with $2n$ first order equations for q_i and p_i.

Example: A Particle in a Potential Field

- The generalized coordinate is simply the position vector $\boldsymbol{r} = (x, y, z)$ and the Lagrangian L is
$$L = \frac{1}{2}m\dot{\boldsymbol{r}}^2 - V(\boldsymbol{r})$$

- The conjugate momenta $p = \frac{\partial L}{\partial \dot{q}}$ are
$$\boldsymbol{p} = m\dot{\boldsymbol{r}}$$
which in this case coincides with the usual definition of momentum.

- The Hamiltonian $H = p_i\dot{q}_i - L$ is
$$H = \boldsymbol{p}\cdot\dot{\boldsymbol{r}} - \left(\frac{1}{2}m\dot{\boldsymbol{r}}^2 - V(\boldsymbol{r})\right) = \frac{\boldsymbol{p}^2}{2m} + V(\boldsymbol{r})$$

Notice the Hamiltonian is written as a function of $\boldsymbol{p}$ in favor of $\dot{\boldsymbol{q}}$.

- The Hamiltonian equations of motion are
$$\dot{\boldsymbol{r}} = \frac{\partial H}{\partial \boldsymbol{p}} = \frac{\boldsymbol{p}}{m} \qquad \text{and} \qquad \dot{\boldsymbol{p}} = -\frac{\partial H}{\partial \boldsymbol{r}} = -\nabla V$$

These equations are familiar; the first is the definition of momentum $\boldsymbol{p} = m\dot{\boldsymbol{r}}$ while the second is Newton's second law of motion for a particle in a potential field.

6.1.4   Hamiltonian Mechanics and Conservation Laws

We return briefly to two of conservation laws and associated constants of motion that we introduced in Lagrangian mechanics. These conservation laws are often simple to see the Hamiltonian formalism.

- If the Hamiltonian H is not explicitly time dependent (i.e. $\frac{\partial H}{\partial t} = 0$) then H is a constant of motion (i.e. $\frac{\mathrm{d}H}{\mathrm{d}t} = 0$).

$$\frac{\mathrm{d}H}{\mathrm{d}t} = \frac{\partial H}{\partial q_i}\dot{q}_i + \frac{\partial H}{\partial p_i}\dot{p}_i + \frac{\partial H}{\partial t} = -p_i\dot{q}_i + \dot{q}_ip_i + \frac{\partial H}{\partial t} = \frac{\partial H}{\partial t}$$

The equality $\frac{\mathrm{d}H}{\mathrm{d}t} = \frac{\partial H}{\partial t}$ means $\frac{\partial H}{\partial t} = 0 \implies \frac{\mathrm{d}H}{\mathrm{d}t} = 0$.

- If an ignorable (aka cyclic) coordinate q doesn't appear in the Lagrangian, the conjugate momentum $p_i = \frac{\partial L}{\partial \dot{q}_i}$ is a constant of motion. Because q doesn't appear in the Lagrangian it certainly doesn't appear in the Hamiltonian, which is constructed from the Lagrangian, so $\frac{\partial H}{\partial q} = 0$. We then have

$$\dot{p}_i \equiv \frac{\partial H}{\partial q_i} = 0$$

so p_q is a constant of motion.

6.1.5   Principle of Least Action à la Hamilton

- In Lagrangian mechanics we defined action as

$$S = \int_{t_1}^{t_2} L(q_i, \dot{q}_i, t) \, \mathrm{d}t = 0$$

and then derived the Lagrange equations of motion from the least action principle by requiring that $\delta S = 0$ for all paths with fixed endpoints so that $\delta q_i(t_1) = \delta q_i(t_2) = 0$.

- In the Hamiltonian formalism, we define action as

$$S = \int_{t_1}^{t_2} (p_i \dot{q}_i - H(q_i, p_i, t))$$

where $\dot{q}_i = \dot{q}_i(q_i, p_i)$. We then consider varying q_i and p_i *independently*, whereas in the Lagrangian formalism a variation of q_i lead to a variation of $\dot{q}_i$. The independent approach makes sense: a theme of the Hamiltonian approach is treating q_i and p_i on equal footing. We have

$$\delta S = \int_{t_1}^{t_2} \left[\delta p_i \dot{q}_i + p_i \delta \dot{q}_i - \frac{\partial H}{\partial q_i} \delta q_i - \frac{\partial H}{\partial p_i} \delta p_i \right] \mathrm{d}t$$

Adding and subtracting $\dot{p}_i \delta q_i$ in the integrand and factoring gives

$$\delta S = \int_{t_1}^{t_2} \left[\left(\dot{q}_i - \frac{\partial H}{\partial p_i} \right) \delta p_i + \left(-\dot{p}_i - \frac{\partial H}{\partial p_i} \right) \delta q_i + \{ p_i \delta \dot{q}_i + \dot{p}_i \delta q_i \} \right] \mathrm{d}t$$

- The last term in the curly brackets can be written as $\frac{\mathrm{d}}{\mathrm{d}t}[p_i \delta q_i]$ and integrated:

$$\delta S = \int_{t_1}^{t_2} \left[\left(\dot{q}_i - \frac{\partial H}{\partial p_i} \right) \delta p_i + \left(-\dot{p}_i - \frac{\partial H}{\partial p_i} \right) \delta q_i \right] \mathrm{d}t + \left[p_i \delta q_i \right]_{t_1}^{t_2}$$

- Hamilton's equations await us in the large parentheses. If we look for extrema $\delta S = 0$ that hold for all δp_i and δq_i we must have

$$\left(\dot{q}_i - \frac{\partial H}{\partial p_i} \right) = 0 \qquad \text{and} \qquad \left(-\dot{p}_i - \frac{\partial H}{\partial p_i} \right) = 0$$

and the boundary conditions $\delta q_i(t_1) = \delta q_i(t_2) = 0$ so that the term $\left[p_i \delta q_i \right]_{t_1}^{t_2}$ vanishes and $\delta S = 0$.

6.2   Charged Particle In an Electromagnetic Field

Solve the equations of motion for a charged particle in an electromagnetic field using both Lagrangian and Hamiltonian formulation of mechanics.

6.2.1   Deriving the Lagrangian

- We start with Maxwell's equations

$$\nabla \times \boldsymbol{E} = -\frac{\partial \boldsymbol{B}}{\partial t} \qquad \text{and} \qquad \nabla \cdot \boldsymbol{B} = 0$$

  Since $\boldsymbol{B}$'s divergence is zero, we can define it in terms of the curl of a vector potential $\boldsymbol{A}$

$$\boldsymbol{B} = \nabla \times \boldsymbol{A}$$

  We substitute $\boldsymbol{B} = \nabla \times \boldsymbol{A}$ into Maxwell's first equation and rearrange to get

$$\nabla \times E + \frac{\partial(\nabla \times \boldsymbol{A})}{\partial t} = \nabla \times \left(\boldsymbol{E} + \frac{\partial \boldsymbol{A}}{\partial t} \right) = 0$$

  Since the curl of the field $\boldsymbol{E} + \frac{\partial \boldsymbol{A}}{\partial t}$ is zero, we can define it terms of the gradient of a scalar field ϕ

$$\boldsymbol{E} + \frac{\partial \boldsymbol{A}}{\partial t} = -\nabla \phi \implies \boldsymbol{E} = -\frac{\partial \boldsymbol{A}}{\partial t} - \nabla \phi$$

  We then substitute $\boldsymbol{E}$ and $\boldsymbol{B}$ into the Lorentz force $\boldsymbol{F} = e(\boldsymbol{E} + \dot{\boldsymbol{r}} \times \boldsymbol{B})$ to get

$$\boldsymbol{F} = e \left(-\frac{\partial \boldsymbol{A}}{\partial t} - \nabla \phi + \dot{\boldsymbol{r}} \times (\nabla \times \boldsymbol{A}) \right)$$

- The next step is to write the force in terms of a velocity-dependent potential $V = V(\boldsymbol{r}, \dot{\boldsymbol{r}})$. Our goal will be to find a potential V satisfying

$$\boldsymbol{F} = \frac{\mathrm{d}}{\mathrm{d}t} \left(\frac{\partial V}{\partial \dot{\boldsymbol{r}}} \right) - \frac{\partial V}{\partial \boldsymbol{r}}$$

  Finding the right V is easier if you first rewrite $\boldsymbol{F}$ using the identities

$$\frac{\partial \boldsymbol{A}}{\partial t} = \frac{\mathrm{d}\boldsymbol{A}}{\mathrm{d}t} - (\dot{\boldsymbol{r}} \cdot \nabla)\boldsymbol{A} \qquad \text{and} \qquad \dot{\boldsymbol{r}} \times (\nabla \times \boldsymbol{A}) = \nabla(\dot{\boldsymbol{r}} \cdot \boldsymbol{A}) - (\dot{\boldsymbol{r}} \cdot \nabla)\boldsymbol{A}$$

  The second identity comes from applying the general vector algebra relation $\boldsymbol{a} \times (\boldsymbol{b} \times \boldsymbol{c}) = \boldsymbol{b}(\boldsymbol{a} \cdot \boldsymbol{c}) - (\boldsymbol{a} \cdot \boldsymbol{b})\boldsymbol{c}$ to $\dot{\boldsymbol{r}} \times (\nabla \times \boldsymbol{A})$, treating ∇ as a vector.

  The first identity comes from rearranging the total derivative

$$\frac{\mathrm{d}}{\mathrm{d}t}(\boldsymbol{A}(\boldsymbol{r}, t)) = \frac{\partial \boldsymbol{A}}{\partial r_i} \dot{r}_i + \frac{\partial \boldsymbol{A}}{\partial t} = (\dot{\boldsymbol{r}} \cdot \nabla)\boldsymbol{A} + \frac{\partial \boldsymbol{A}}{\partial t}$$

  Plugging the two expressions into the Lorentz force gives

$$\boldsymbol{F} = e \left(-\nabla \phi - \frac{\mathrm{d}A}{\mathrm{d}t} + (\dot{\boldsymbol{r}} \cdot \nabla)\boldsymbol{A} + \nabla(\dot{\boldsymbol{r}} \cdot \boldsymbol{A}) - (\dot{\boldsymbol{r}} \cdot \nabla)\boldsymbol{A} \right)$$

$$= e \left(-\nabla \phi - \frac{\mathrm{d}A}{\mathrm{d}t} + \nabla(\dot{\boldsymbol{r}} \cdot \boldsymbol{A}) \right)$$

69

- It takes a bit of reverse-engineering, but it turns out the potential

$$V(\boldsymbol{r}, \dot{\boldsymbol{r}}) = e(\phi - \dot{\boldsymbol{r}} \cdot \boldsymbol{A})$$

satisfies our desired expression $\boldsymbol{F} = \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial V}{\partial \dot{\boldsymbol{r}}}\right) - \frac{\partial V}{\partial \boldsymbol{r}}$. Here's a quick confirmation

$$\frac{\partial V}{\partial \boldsymbol{r}} = e\left(\frac{\partial \phi}{\partial \boldsymbol{r}} - \frac{\partial}{\partial \boldsymbol{r}}(\dot{\boldsymbol{r}} \cdot \boldsymbol{A})\right) = e\left(\nabla\phi - \frac{\partial \dot{\boldsymbol{r}}}{\partial \boldsymbol{r}}\boldsymbol{A} - \frac{\partial \boldsymbol{A}}{\partial r_i}\dot{r}_i\right)$$
$$= e\left(\nabla\phi - 0 \cdot \boldsymbol{A} - \nabla(\dot{\boldsymbol{r}} \cdot \boldsymbol{A})\right) = e\nabla\phi - e\nabla(\dot{\boldsymbol{r}} \cdot \boldsymbol{A})$$

and

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial V}{\partial \dot{\boldsymbol{r}}}\right) = e\frac{\mathrm{d}}{\mathrm{d}t}\left[\frac{\partial \phi}{\partial \dot{\boldsymbol{r}}} - \frac{\partial \dot{\boldsymbol{r}} \cdot \boldsymbol{A}}{\partial \dot{\boldsymbol{r}}}\right] = e\frac{\mathrm{d}}{\mathrm{d}t}(0 - \boldsymbol{A}) = -e\frac{\mathrm{d}\boldsymbol{A}}{\mathrm{d}t}$$

Combining these and rearranging gives

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial V}{\partial \dot{\boldsymbol{r}}}\right) - \frac{\partial V}{\partial \boldsymbol{r}} = -e\frac{\mathrm{d}\boldsymbol{A}}{\mathrm{d}t} - e\nabla\phi + e\nabla(\dot{\boldsymbol{r}} \cdot \boldsymbol{A}) = \boldsymbol{F}$$

Having found the correct potential $V(\boldsymbol{r}, \dot{\boldsymbol{r}})$, the Lagrangian is

$$L = T - V = \frac{m\dot{\boldsymbol{r}}^2}{2} - e(\phi - \dot{\boldsymbol{r}} \cdot \boldsymbol{A})$$

6.2.2   Solution with Lagrangian Mechanics

- The Lagrangian for a particle of charge e in an electromagnetic field is

$$L = \frac{1}{2}m\dot{\boldsymbol{r}}^2 - e(\phi - \dot{\boldsymbol{r}} \cdot \boldsymbol{A})$$

Note the Lagrangian splits into the usual kinetic energy term $\frac{1}{2}m\dot{\boldsymbol{r}}^2$ and an additional term due to the EM interaction. The momentum $\boldsymbol{p}$ conjugate to the coordinate $\boldsymbol{r}$ is

$$\boldsymbol{p} = \frac{\partial L}{\partial \dot{\boldsymbol{r}}} = m\dot{\boldsymbol{r}} + e\boldsymbol{A}$$

Note that the momentum is not simply the Newtonian $m\dot{\boldsymbol{r}}$, but is modified in the present of an EM field.

- With Lagrangian in hand, we can calculate Lagrange's equations

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{\boldsymbol{r}}}\right) - \frac{\partial L}{\partial \boldsymbol{r}} = \frac{\mathrm{d}}{\mathrm{d}t}(m\dot{\boldsymbol{r}} + e\boldsymbol{A}) + e\nabla\phi - e\nabla(\dot{\boldsymbol{r}} \cdot \boldsymbol{A})$$

In terms of components and the indices $a, b \in \{1, 2, 3\}$ of the Cartesian components, the equation of motion reads

$$m\ddot{r}^a = -e\left(\frac{\partial \phi}{\partial r^a} + \frac{\partial A_a}{\partial t}\right) + e\left(\frac{\partial A_b}{\partial r^a} - \frac{\partial A_a}{\partial r^b}\right)\dot{r}^b$$

- In terms of indices and components, the definitions of the $\boldsymbol{E}$ and $\boldsymbol{B}$ fields are

$$E_a = -\frac{\partial \phi}{\partial r^a} - \frac{\partial A_a}{\partial t} \qquad \text{and} \qquad B_a = \epsilon_{cab}\frac{\partial A_b}{\partial r^a} \iff \epsilon_{abc}B_c = \frac{\partial A_b}{\partial r^a} - \frac{\partial A_a}{\partial r^b}$$

But these expressions agree exactly with the terms in parentheses in the equation of motion! Substitution, the equation of motion becomes

$$m\ddot{r}^a = eE_a + \epsilon_{abc}e\dot{r}^b B_c \qquad \text{or} \qquad m\ddot{\boldsymbol{r}} = e(\boldsymbol{E} + \dot{\boldsymbol{r}} \times \boldsymbol{B})$$

which is the Lorentz force law!

6.2.3   Solution with Hamiltonian Mechanics

- The Lagrangian for a particle of charge e in an EM field is

$$L = \frac{1}{2}m\dot{\boldsymbol{r}}^2 - e(\phi - \dot{\boldsymbol{r}} \cdot \boldsymbol{A})$$

  while the momentum $\boldsymbol{p}$ conjugate to the position $\boldsymbol{r}$ is

$$\boldsymbol{p} = \frac{\partial L}{\partial \dot{\boldsymbol{r}}} = m\dot{\boldsymbol{r}} + e\boldsymbol{A}$$

- Inverting the expression for $\boldsymbol{p}$ gives us

$$\dot{\boldsymbol{r}} = \frac{1}{m}(\boldsymbol{p} - e\boldsymbol{A})$$

  from which we can calculate the Hamiltonian

$$H(\boldsymbol{p}, \boldsymbol{r}) = \boldsymbol{p}\dot{\boldsymbol{r}} - L = \frac{\boldsymbol{p}}{m}(\boldsymbol{p} - e\boldsymbol{A}) - \left[\frac{1}{2m}(\boldsymbol{p} - e\boldsymbol{A})^2 - e\phi + \frac{e}{m}(\boldsymbol{p} - e\boldsymbol{A}) \cdot \boldsymbol{A}\right]$$

$$= \frac{1}{2m}(\boldsymbol{p}^2 - 2e\boldsymbol{p} \cdot \boldsymbol{A} + e^2\boldsymbol{A}^2) + e\phi$$

$$= \frac{1}{2m}(\boldsymbol{p} - e\boldsymbol{A})^2 + e\phi$$

- The first equation of motion is

$$\dot{\boldsymbol{r}} = \frac{\partial H}{\partial \boldsymbol{p}} = \frac{1}{m}(\boldsymbol{p} - e\boldsymbol{A})$$

  The equation for $\dot{\boldsymbol{p}} = -\frac{\partial H}{\partial \boldsymbol{r}}$ is best given in terms of indices and components

$$\dot{p}_a = -\frac{\partial H}{\partial r_a} = \frac{e}{m}(p_b - eA_b)\frac{\partial A_b}{\partial r_a} - e\frac{\partial \phi}{\partial r_a}$$

  some manipulation of the indices and the definition of the cross product in component form shows this is equivalent to the Lorentz force $\boldsymbol{F} = e(\boldsymbol{E} - \boldsymbol{v} \times \boldsymbol{B})$.

Concrete Example: Consider a particle in a uniform one-directional magnetic field $\boldsymbol{B} = (0, 0, B)$.

- The magnetic field $\boldsymbol{B} = (0, 0, B)$ corresponds to a vector potential

$$\boldsymbol{A} = (-By, 0, 0)$$

  Consider a particle moving in the (x, y) plane. Then $\boldsymbol{p} = (p_x, p_y, 0)$ and the Hamiltonian is

$$H = \frac{1}{2m}\left[p_x^2 + p_y^2 + 2ep_xBy + e^2B^2y^2\right]$$

$$= \frac{1}{2m}(p_x + eBy)^2 + \frac{p_y^2}{2m}$$

- The equations of motion are

$$\dot{x} = \frac{\partial H}{\partial p_x} = \frac{1}{m}(p_x + eBy) \qquad \dot{y} = \frac{\partial H}{\partial p_y} = \frac{p_y}{m} \qquad \dot{z} = 0$$

$$\dot{p}_y = -\frac{\partial H}{\partial y} = -\frac{eB}{m}(p_x + eBy) \qquad \dot{p}_x = \dot{p}_z = 0$$

- Next, we note that

$$eBx + m\dot{y} = \text{constant} \equiv a \qquad \text{and} \qquad p_x = m\dot{x} - eBy = \text{constant} \equiv b$$

  The first equality can be found by combining the equations for $\dot{x}$ and $\dot{p}_y$, integrating and using $p_y = m\dot{y}$; the second e.g. by definition $p_x = \frac{\partial L}{\partial \dot{x}}$ from the Lagrangian.

- The result is a simple system of coupled LDEs for x and y

$$m\dot{y} + eBx = a \qquad \text{and} \qquad m\dot{x} - eBy = b$$

  The system is essentially a harmonic oscillator; the solution is sinusoidal

$$x = \frac{a}{eB} + R\sin(\omega(t - t_0)) \qquad \text{and} \qquad y = -\frac{b}{eB} + R\cos(\omega(t - t_0))$$

  where a, b, R and t_0 determined by the initial conditions. The particle makes circles of radius R in the (x, y) plane at the *cyclotron frequency*

$$\omega = \frac{eB}{m}$$

6.3   Poisson Brackets

What are Poisson properties? State their basic properties and discuss, with examples, their role in Hamiltonian mechanics.

- Let $f(q_i, p_i)$ and $g(q_i, p_i)$ be two functions defined on phase space where $i = 1, \ldots, N$. The *Poisson bracket* of the functions f and g is

$$\{f, g\} = \sum_{i=1}^{N} \left(\frac{\partial f}{\partial q_i} \frac{\partial g}{\partial p_i} - \frac{\partial f}{\partial p_i} \frac{\partial g}{\partial q_i} \right) \equiv \frac{\partial f}{\partial q_i} \frac{\partial g}{\partial p_i} - \frac{\partial f}{\partial p_i} \frac{\partial g}{\partial q_i}$$

  We generally use Einstein summation notation and drop the sum symbol $\sum$.

- Some properties of Poisson brackets are:

  - $\{f, g\} = -\{g, f\}$
  - Linearity: $\{\alpha f + \beta g, h\} = \alpha\{f, h\} + \beta\{g, h\}$ for all real scalars $\alpha, \beta \in \mathbb{R}$
  - Leibniz (product) rule: $\{fg, h\} = f\{g, h\} + g\{f, h\}$
  - Jacobi identity: $\{f, \{g, h\}\} + \{h, \{f, g\}\} = 0$

  Note that the Poisson bracket obeys a similar algebraic structure as the differentiation operator d and the matrix commutator $[\cdot, \cdot]$.

- The next set of identities is

$$\{q_i, q_j\} = 0 \qquad \{p_i, p_j\} = 0 \qquad \{q_i, p_j\} = \delta_{ij}$$

  All three are proved straightforwardly with the trivial relations

$$\frac{\partial p_i}{\partial q_j} = \frac{\partial q_i}{\partial p_j} = 0 \qquad \text{and} \qquad \frac{\partial p_i}{\partial p_j} = \frac{\partial q_i}{\partial q_j} = \delta_{ij}$$

- Next, a more interesting identity: for any function $f(q_i, p_i, t)$ we have

$$\frac{\mathrm{d}f}{\mathrm{d}t} = \{f, H\} + \frac{\partial f}{\partial t}$$

We first differentiate f, then recognize the Hamilton equations $\dot{q}_i = \frac{\mathrm{d}q}{\mathrm{d}t} = \frac{\partial H}{\partial p_i}$ and $\dot{p} = \frac{\mathrm{d}p}{\mathrm{d}t} = -\frac{\partial H}{\partial q}$

$$\frac{\mathrm{d}f}{\mathrm{d}t} = \frac{\partial f}{\partial q_i}\frac{\mathrm{d}q_i}{\mathrm{d}t} + \frac{\partial f}{\partial p_i}\frac{\mathrm{d}p_i}{\mathrm{d}t} + \frac{\partial f}{\partial t}\frac{\mathrm{d}t}{\mathrm{d}t} = \frac{\partial f}{\partial q_i}\frac{\partial H}{\partial p_i} - \frac{\partial f}{\partial p_i}\frac{\partial H}{\partial q_i} + \frac{\partial f}{\partial t}$$

$$= \{f, H\} + \frac{\partial f}{\partial t}$$

The result is important. It tells us that any function $I = I(p_i, q_i)$ for which $\{I, H\} = 0$ is a constant of motion: if $\{I, H\} = 0$ then $\frac{\mathrm{d}f}{\mathrm{d}t} = \frac{\partial f}{\partial t}$ and $\frac{\partial f}{\partial t} = 0 \implies \frac{\mathrm{d}f}{\mathrm{d}t} = 0$. If $\{I, H\} = 0$ we say that the functions I and H *Poisson commute*.

As an example, we know from Lagrangian mechanics that if q_i is a cyclic coordinate (i.e. $\frac{\partial H}{\partial q_i} = 0$) then the corresponding generalized momentum p_i is a constant of motion. This also holds in language of Poisson brackets: using the identities $\frac{\partial p}{\partial q} = 0$ and $\frac{\partial p_i}{\partial p_j} = \delta_{ij}$ and then applying $\dot{p}_i \equiv 0$ we have

$$\{p_i, H\} = \frac{\partial p_i}{\partial q_j}\frac{\partial H}{\partial p_j} - \frac{\partial p_i}{\partial p_j}\frac{\partial H}{\partial q_j} = 0 - \frac{\partial H}{\partial q_i} = \dot{p}_i = 0$$

- Next, an important consequence of the Jacobi identity and $\{f, g\} = -\{g, f\}$. If I and J are both constants of motion, then

$$\{\{I, J\}, H\} = \{I, \{J, H\}\} + \{\{I, H\}, J\} = 0$$

which means $\{I, J\}$ is a constant of motion. In other words, if we know two constants of motion, we can generate more using the Poisson brackets and the Jacobi identity.

Poisson Brackets and Angular Momentum

- Poisson bracket show us something cool: if two components of angular momentum (e.g. L_1 and L_2) are conserved, the third component L_3 must also be conserved. This is an application of the theorem stating $\{I, J\}$ is conserved if I and J are separately conserved. We just need to show that $\{L_1, L_2\} = L_3$. We have

$$L_1 = r_2 p_3 - r_3 p_2 \qquad \text{and} \qquad L_2 = r_3 p_1 - r_1 p_3$$

$$\frac{\partial L_1}{\partial \boldsymbol{r}} = (0, p_3, -p_2) \qquad \text{and} \qquad \frac{\partial L_2}{\partial \boldsymbol{r}} = (-p_3, 0, p_1)$$

$$\frac{\partial L_1}{\partial \boldsymbol{p}} = (0, -r_3, r_2) \qquad \text{and} \qquad \frac{\partial L_2}{\partial \boldsymbol{p}} = (r_3, 0, -r_1)$$

Now we just need to put the pieces together. Using the dot product gives

$$\{L_1, L_2\} = \frac{\partial L_1}{\partial \boldsymbol{r}} \cdot \frac{\partial L_2}{\partial \boldsymbol{p}} - \frac{\partial L_1}{\partial \boldsymbol{p}} \cdot \frac{\partial L_2}{\partial \boldsymbol{r}} = r_1 p_2 - r_2 p_1 = L_3$$

- Now if L_1, L_2 and L_3 are conserved, the entire vector $\boldsymbol{L}$ is conserved. In other words, if two components (e.g. L_1 and L_2) are conserved, the entire vector $\boldsymbol{L}$ is automatically conserved. Note also the quantum mechanic-like relationship

$$\{L^2, L_i\} = \{L_1^2 + L_2^2 + L_3^2, L_i\} = 0 \qquad \text{for} \qquad i = 1, 2, 3$$

73