Preface

Topics to be covered in this course:

1. Lagrangian Mechanics and Applications
2. Rigid body Motion
3. Hamiltonian Dynamics

1. Lecture 1 - Introduction to Analytical Mechanics

There is a magic trick: Imagine a mass and a spring, with no gravity, with 0 stretched length. The force on the spring is simply $F = -kx$, which is simple. We can solve this using $\vec{F} = m\vec{a}$, but instead let us attempt something else, writing the following linear combination of energies.

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

Where T represents the kinetic energy of the mass, V is the potential energy of the mass, and $\mathcal{L}$ is commonly referred to as the *Lagrangian*. In this case, $T = \frac{1}{2}m\dot{x}^2$, where $\dot{x} = \mathrm{d}x/\mathrm{d}t$. The potential energy would then be $V = \frac{1}{2}kx^2$. Therefore, the Lagrangian is:

$$\mathcal{L} = \mathcal{L}(x, \dot{x}) = \frac{1}{2}m\dot{x}^2 - \frac{1}{2}kx^2$$

Using the Euler-Lagrange equation (ignore where this comes from for now),

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial \mathcal{L}}{\partial \dot{x}}\right) = \frac{\partial \mathcal{L}}{\partial x} \tag{1}$$

Which allows us to optimise for the stationary points of a given functional. Substituting the appropriate expressions,

$$\Rightarrow m\ddot{x} = -kx$$

Which is equivalent to the initial expression for $\vec{F} = m\vec{a}$ for our system. This works for many other problems. In general, given some arbitrary potential $V = V(x)$, we see $m\ddot{x} = -\frac{\partial V}{\partial x}$ which is equivalent to the force (conservative forces), where Newton's 2nd law is recovered.

For a three dimensional system, $x_1 = x, x_2, = y, x_3 = z$, we write:

$$T = \frac{1}{2}m\sum_i \dot{x_i}^2 \tag{2}$$

$$\Rightarrow \mathcal{L} = \frac{1}{2}m\sum_i \dot{x_i}^2 - V(x_i) \tag{3}$$

To solve for the motion of the particle now, we then have Euler-Lagrange equations to be solved (details in "Calculus of Variations" by Fomin), one for each coordinate. These essentially form the components of a vector equation.

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial \mathcal{L}}{\partial \dot{x_i}}\right) = \frac{\partial \mathcal{L}}{\partial x} \tag{4}$$

Equivalent to:

$$m\ddot{\vec{r}} = -\vec{\nabla} V(\vec{r}) \tag{5}$$

$$m\vec{a} = \vec{F} \tag{6}$$

Where $\vec{r}$ is shorthand for the position coordinates.

Note that we don't necessarily always want to work in Cartesian coordinates. Imagine a pendulum mass system where the string is constructed out of a spring with equilibrium length l, spring constant k. The coordinates we want to use are θ, x - does the Euler-Lagrange equation still hold true in exactly the same way? Let us verify.

We still have $T = \frac{1}{2}m\vec{v}^2$ - but what is $\vec{v}^2$? It contains a radial component and an angular component. We define the square of the vector by its inner product. In this case, $\vec{v}^2 = \vec{v}_r^2 + \vec{v}_\theta^2$, and $\vec{v}_\theta^2 = (l+x)^2\dot{\theta}^2$.

$$\Rightarrow T = \frac{1}{2}m(\dot{x}^2 + (l+x)^2\dot{\theta}^2) \tag{7}$$

What about the potential energy V of the system? There is the spring constant, and there is also gravity. We write:

$$V = \frac{1}{2}kx^2 - mg(l+x)\cos\theta \tag{8}$$

We now need to re-write the Euler-Lagrange equation in our new coordinate system. For x:

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial \mathcal{L}}{\partial \dot{x}}\right) = \frac{\partial \mathcal{L}}{\partial x} \tag{9}$$

This provides us with:

$$m\ddot{x} = m(l+x)\dot{\theta}^2 + mg\cos\theta - kx \tag{10}$$

For θ, we have:

$$\frac{\mathrm{d}}{\mathrm{d}t}\left[m(l+x)^2\dot{\theta}\right] = -mg(l+x)\sin\theta \tag{11}$$

$$\Rightarrow m(l+x)\ddot{\theta} + 2m\dot{x}\theta' = -mg\sin\theta \tag{12}$$

In equation (10), we actually have a centripetal acceleration term on the right hand side. In general, the radial component of $\vec{a}$ we have:

$$\ddot{x} - \frac{V_\theta^2}{r} = \ddot{x} - (l+x)\ddot{\theta}^2 \tag{13}$$

The other equation tells us that torque is the momentum of the angular rotation.

2. Lecture 2 - Generalized Coordinates and Principle of Least Action

We have that the Euler-Lagrange equation is equivalent to Newton's second law. From the previous example, we have the polar coordinate system (x, θ) giving us the two equations:

$$q_1 = x,\ m\ddot{x} - m(l+x)\dot{\theta}^2 = mg\cos\theta - kx \tag{14}$$

$$q_2 = \theta,\ \frac{\mathrm{d}}{\mathrm{d}t}\left[m(l+x)^2\dot{\theta}\right] = -mg(l+x)\sin\theta \tag{15}$$

The LHS of the first equation is the acceleration in the radial direction.

The second equation above indicates to us that the torque is the rate of change of the angular momentum. The term in the square brackets is equal to $I\omega$, where I is the moment of inertia. On the right hand side we actually have $\vec{r} \times \vec{F}$, the torque. We actually don't even have to know about moment of inertia or torque for these terms to show up.

In the Lagrangian scheme, we don't have to actually discuss about forces. Instead, we are talking about energies and potentials. We are plugging and chugging, and the $\vec{F} = m\vec{a}$ equations automatically pop out. Note that the Euler-Lagrange equation is acutally invariant under different coordinate systems.

Let us define some terminology: We refer to "equations of motions" as the equations that show up from solving the Euler-Lagrange equations. We are intersted in the motion of point-masses for now, which means that we only care about the position of the particle for now (we will consider how the particle rotates during its motion later on), allowing us to describe the particle's position using some function $\vec{r}(t)$.

In Cartesian coordinates, $\vec{r}(t) = x(t)\hat{x}+y(t)\hat{y}+z(t)\hat{z}$. Cartesian coordinates might look gnarly when dealing with problems containing spherical symmetry, which prompts one to ask: What is the best coordinate system to use? One can use spherical coordinates, with $\vec{r}(t) = r(t)\hat{r}+\theta(t)\hat{\theta}+\varphi(t)\hat{\varphi}$ parametrised differently.

Lastly, we have a cylindrical coordinate, parametrized by $\vec{r}(t) = \rho(t)\hat{r} + \varphi(t)\hat{\varphi} + z(t)\hat{z}$ for cylindrical symmetry. These three coordinate systems are the most common coordinate systems to be using, and in this course we mainly be dealing with these three coordinates.

The nice thing about Lagrangian mechanics is that we do not have to choose an inertial frame of reference when dealing with rigid body motion.

To specify the position of a particle in 3D, we will need 3 coordinates $\{q_i\}$. Generalizing this, for a system of n-particles in 3D we will need $3N$ specific coordinates to describe the whole system, $\{q_i(t)\}, i = 1, ..., 3N$. These q_i are referred to as "generalized coordinates", as we can express the coordinates in whichever system we choose. Likewise for velocity, we have generalized velocity which is simply the time derivative of each q_i.

Definition: The number of generalized coordinates required to describe a system fully is known as the *degrees of freedom* of the system.

The idea is that the Euler-Lagrange equation will provide a differential equation, equivalent to $\vec{F} = m\vec{a}$ to be solved, giving the equations of motion.

Now we will discuss the Principle of Least Action. Let us consider a single system with 1 degree of freedom for simplicity. At time t_0, we know the particle is at $q(t_0)$. Suppose we know that at time t_1, we know it is

at $q(t_1)$. The question we ask is: What path did the particle take? There are an infinite number of possible paths to take in principle - but physically there exists only one path it takes. This is due to the principle of least action.

For each path, we assign a number (essentially, we design a functional for the paths). This functional is known as the action, S defined as:

$$S = \int_{t_0}^{t_1} \mathcal{L}(q(t), \dot{q}(t), t)\mathrm{d}t \tag{16}$$

We now make an extraordinary claim: The particle takes the path which minimizes the functional (technically it is for which the functional is stationary). This is the principle of least action - whilst this seems completely arbitrary and random, it works.

Note that S has units of $\mathrm{energy} \times \mathrm{time}$. We allow $\mathcal{L}$ to have explicit time dependence. However, S cannot be a function dependent on t, as we are integrating with respect to time. S is a function dependent on paths, which is infinite, so we actually have an infinite number of variables to consider for S. Along each path, it depends on $q(t), \dot{q}(t)$ at every point along the path. Hence, $S = S(q(t_0), \dot{q}(t_0), q(t_0+\varepsilon), \dot{q}(t_0+\varepsilon), ...)$ in the limit where $\varepsilon \to 0$.

We are optimizing for the stationary points of a functional, $S[q(t)]$. Recall that for a function of 1 variable, around the neighbourhood of each extremum the sign of the gradient is the same. This implies that $f(x_0 + \delta) - f(x_0) = O(\delta^2)$.

For functionals, consider small variations in the path: $q_c(t) \to q_c(t) + \delta q(t)$, but with the fixed endpoints $\delta q(t_0) = \delta q(t_1) = 0$. Then, if $q_c(t)$ is an extremum of S, we have that:

$$S[q_c(t) + \delta q(t)] - S[q_c(t)] \approx O(\delta q^2) \tag{17}$$

For now we accept the fact that the principle of stationary action is taken as granted, and that it is as "fundamental" as it gets. It is a mathematical property to the solutions to $\vec{F} = m\vec{a}$.

3. Lecture 3 - Derivation of Euler-Lagrange Equation

Imagine varying a path slightly $\beta(t)$ from the one of least action, with the property of fixed endpoints. If the action is an extremum, then its derivative with respect to a must be 0.

$$\Rightarrow S[q_c(t) + a\beta(t)] \tag{18}$$

$$\Rightarrow \frac{\partial}{\partial a} S[q_a(t)] = 0 \tag{19}$$

At $a = 0$ for extremum. Let $q_a(t) = a\beta(t)$,

$$\frac{\partial}{\partial a}S[q_a(t)] = \frac{\partial}{\partial a}\int_{t_0}^{t_1}\mathcal{L}(q_a(t),\dot{q}_a(t),t)\mathrm{d}t \tag{20}$$

$$\text{moving the derivative inside } = \int_{t_0}^{t_1}\frac{\partial}{\partial a}\mathcal{L}(q_a(t),\dot{q_a}(t),t)\mathrm{d}t \tag{21}$$

$$\text{using the chain rule } = \int_{t_0}^{t_1}\left(\frac{\partial\mathcal{L}}{\partial q_a}\frac{\partial q_a}{\partial a}+\frac{\partial\mathcal{L}}{\partial\dot{q_a}}\frac{\partial\dot{q_a}}{\partial a}\right)\mathrm{d}t \tag{22}$$

$$\text{simplifying } = \int_{t_0}^{t_1}\left[\frac{\partial\mathcal{L}}{\partial q_a}\beta(t)+\frac{\partial\mathcal{L}}{\partial\dot{q}}\dot{\beta}(t)\right]\mathrm{d}t \tag{23}$$

However, using integration by parts (and the fact that the endpoints are fixed),

$$\int_{t_0}^{t_1}\frac{\partial\mathcal{L}}{\partial\dot{q_a}}\dot{\beta}(t)\mathrm{d}t = \frac{\partial\mathcal{L}}{\partial\dot{q_a}}\beta(t)|_{t_0}^{t_1}-\int_{t_0}^{t_1}\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial\mathcal{L}}{\partial\dot{q_a}}\right)\beta(t)\mathrm{d}t \tag{24}$$

Which implies that:

$$\frac{\partial}{\partial a}S[q_a(t)] = \int_{t_0}^{t_1}\left(\frac{\partial\mathcal{L}}{\partial q_a}-\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial\mathcal{L}}{\partial\dot{q_a}}\right)\right)\beta(t)\mathrm{d}t \tag{25}$$

By letting $a \to 0$, we have that $q_a(t) \to q_c(t)$. Consider that $\beta(t)$ can be ANY function, giving us the equation:

$$\Rightarrow\int_{t_0}^{t_1}\left[\frac{\partial\mathcal{L}}{\partial q_c}-\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial\dot{q_c}}\right)\right]\beta(t)\mathrm{d}t = 0 \tag{26}$$

This implies that the term inside the square brackets must be equal to 0. Hence,

$$\therefore\frac{\partial\mathcal{L}}{\partial q_c}-\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial\mathcal{L}}{\partial\dot{q_c}}\right) = 0 \tag{27}$$

Giving us the Euler-Lagrange equation. Note that in this derivation, we never restricted the integrand to being the Lagrangian $\mathcal{L}$. This is much more powerful, and in general works for any functional (and we can also generalize this to any number of variables).

What if we have N-degrees of freedom? (i.e. $q_i(t), i = 1, ..., N$) The beauty of this is that the Euler-Lagrange equation remains the same, and we simply solve for a system of N variables. Any change of coordinate systems will have the same Euler-Lagrange equation.

Note: For a given system, the Lagrangian $\mathcal{L}$ is not unique.

- We can always multiply the Lagrangian by a constant α, and it should still give us the same equations of motion.
- We can add a total time derivative of any function. $\mathcal{L} \to \mathcal{L} + \frac{\mathrm{d}f}{\mathrm{d}t}$. Why? Consider $\Delta S = \int_{t_0}^{t_1}\frac{\mathrm{d}f}{\mathrm{d}t}\mathrm{d}t = f(q_i(t_1),t_1) - f(q_i(t_0),t_0)$, and the endpoints are the same for all paths. Hence, it doesn't vary the stationary points of the action.

4. Lecture 4 - Origins of the Lagrangian

Why is $\mathcal{L} = T - V$? Let us assume a free particle, so there is no potential. Landau and Lifschitz showed that then $\mathcal{L} \propto |\vec{v}|^2$. Since we can scale the Lagrangian by any constant of choice, we can scale it to $\frac{1}{2}m\vec{v}^2$.

Where does the Principle of Least Action come from? It is from taking the classical limit of quantum mechanics. The units of the action, S are in $E \times$ time, which has the exact same units of Planck's constant, $\hbar \approx 10^{-34} Js$.

We can obtain a dimensionless quantity $S/\hbar$. If $S/\hbar >> 1$, then we are in a classical world, but if $S/\hbar \leq 1$, we are dealing with quantum mechanics. By taking the limit of the double slit experiment to an infinite number of slits, we find that the amplitude for a particle to go from A to B is the sum over all paths from A to B.

$$\sum_{\text{All paths from A to B}} \text{Amp(Path)} \tag{28}$$

Feynman argued that the amplitude of a path is given by:

$$\text{Amp(path)} = e^{-i\frac{S_{\text{path}}}{\hbar}} \tag{29}$$

Which turns out to be equivalent to Schroedinger's equation. How does this system behave when $S >> \hbar$? We are essentially considering a classical system. Imagine a scenario with the same endpoints, but two paths very close to each other.

Let us have $S, S + \delta S$ as our two paths. If $S, \delta S >> \hbar$, then two similar paths will have wildly different phases. If we sum up over all possible nearby paths, the amplitudes of the paths would destructively interfere.

However, there is one remaining possible path: The path of stationary action, where the phases would constructively interfere, as nearby paths would have $\delta S \approx 0$. This is the classical path we obtain.

Constrained Systems

"Holonomic constraints" - they are things that can be written in the following form:

$$f_i(q_1, ..., q_N) = 0 \tag{30}$$

We can then solve for this equation, and eliminate one of the degrees of freedom.

5. Lecture 5 - Constrained Systems

For example, two blocks with a fixed length between them is a Holonomic constraint, as the distance between the two blocks can be expressed as $x_1 - x_2 = L$. This implies we can express one variable as another, and reduce the degrees of freedom by one.

Mathematically, this essentially implies that our system of equations are linearly dependent. Hence, one of the coordinates can be represented as a linear combination of the others. Obviously, we are implying that

f is a linear combination of the other variables for this explanation to work, but we will skip out the finer details for now.

Another example is the simple pendulum. Using $\vec{F} = m\vec{a}$, we identify the forces present in the system, and solve for the system. Using the Lagrangian and Holonomic constraints, we use the fact that $x^2 + y^2 = l^2$ giving us the two equations:

$$\ddot{\theta} = \frac{-g}{l} \sin\theta \tag{31}$$

$$T = ml\dot{\theta}^2 + mg\cos\theta \tag{32}$$

Typically, we can choose a generalized coordinate system where we have already implemented our constraints. In the case of the simple pendulum, we can simply choose our system to have 1 coordinate only: The θ component, as the length of the string is fixed. If we used the Euler-Lagrange equation again, and we have:

$$\ddot{\theta} = \frac{-g}{l} \sin\theta \tag{33}$$

Where we only have one equation this time - same equation, but less amount of work.

Spherical Pendulum

Consider a 3D-pendulum, which can move in directions perpendicular to the plane of oscillation. It will have two degrees of freedom: φ and θ. We need to express $\vec{v}^2$ in terms of $\theta, \dot{\theta}, \varphi, \dot{\varphi}$. Since we know that:

$$x = r\sin\theta\cos\varphi \tag{34}$$

$$y = r\sin\theta\sin\varphi \tag{35}$$

$$z = r\cos\theta \tag{36}$$

This gives us that:

$$\vec{v}^2 = \dot{x}^2 + \dot{y}^2 + \dot{z}^2 \tag{37}$$

$$= \dot{r}^2 + r^2\dot{\theta}^2 + r^2\dot{\varphi}^2\sin^2\theta \tag{38}$$

Then we have the Lagrangian:

$$\mathcal{L} = \frac{1}{2}ml^2(\dot{\theta}^2 + \dot{\varphi}^2\sin^2\theta) + mgl\cos\theta \tag{39}$$

This provides us with two Euler-Lagrange equations. With respect to θ, it gives us some equation (which we don't care so much in this context). However, with respect to φ, we find that:

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial\mathcal{L}}{\partial\dot{\varphi}}\right) = 0 \tag{40}$$

$$\Rightarrow \frac{\mathrm{d}}{\mathrm{d}t}(ml^2\sin^2\theta\dot{\varphi}) = 0 \tag{41}$$

This is a conservation law, where the angular momentum is conserved as our pendulum is oscillating around the z-axis.

Pendulum and Rotating Disc

Imagine a pendulum attached to a rotating disc with angular velocity ω, a pendulum attached to a fixed disc, a pendulum attached to a disc free to spin in whichever way it wants.

They will have $1, 1, 2$ degrees of freedom respectively. Consider the first scenario. How do we find the velocity?

We define the position of the pivot to be (where the origin is the center of the disc):

$$x_p = R\cos(\omega t) \tag{42}$$
$$y_p = R\sin(\omega t) \tag{43}$$

Then we find that the x, y of the pendulum can be expressed as:

$$x = x_p + l\sin\theta = R\cos\omega t + l\sin\theta \tag{44}$$
$$y = y_p - l\cos\theta = -R\sin\omega t - l\cos\theta \tag{45}$$

We differentiate it with respect to time, and put it into the Lagrangian:

$$\mathcal{L} = \frac{1}{2}m(R^2\omega^2 - 2l\dot{\theta}R\omega\cos(\theta+\omega t) + l^2\dot{\theta}^2) + mg(R\sin\omega t + l\sin\theta) \tag{46}$$

Applying EL gives us:

$$\ddot{\theta} + \frac{g}{l}\sin\theta = \frac{R\omega^2}{l}\cos(\theta+\omega t) \tag{47}$$

This precisely gives us the equations of motion, which would be slightly non-trivial in the image of $\vec{F} = m\vec{a}$.

6. Lecture 6 - Conservation Laws I

If $\frac{\partial\mathcal{L}}{\partial q_0} = 0$, then q_i is "cyclic". If $\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial\mathcal{L}}{\partial q_i}\right) = 0$, then $\frac{\partial\mathcal{L}}{\partial q_i}$ is conserved. Whilst the Lagrangian might not help us solve the equations of motion, it tells us other characteristics of the system that does not arise from $\vec{F} = m\vec{a}$.

Going back to the example of the spherical pendulum, we have eq. 41 giving us that $(\vec{r}\times\vec{p})_z$, the $\hat{z}$ component of angular momentum, indicating the conservation of angular momentum in the example.

Consider a free particle in 3D

Let us use the Cartesian coordinates. Then,

$$\mathcal{L} = \frac{1}{2}m\sum_\mu \dot{x}_\mu^2 \rightarrow \text{all } x_i's \text{ are cyclic} \tag{48}$$

Since there is no x_μ dependence,

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial\dot{x}_i}\right) = 0 \tag{49}$$
$$\Rightarrow \frac{\partial\mathcal{L}}{\partial\dot{x}_i} = m\dot{x}_i \text{ is conserved} \tag{50}$$

Where the term of the right is the linear momentum. Now we want to generalize this, so let us define the generalized momentum:

$$p_i = \frac{\partial \mathcal{L}}{\partial q_i} \tag{51}$$

Re-visiting the spherical coordinate, since $V(x, y, z) = V(z)$, then $p_1 = \dot{x_1}$ and $p_2 = \dot{x_2}$ are conserved (but not p_3). We also found that:

$$p_\varphi = ml^2 \sin^2\theta\varphi^2 \text{ is conserved} \tag{52}$$

Theorem: Everytime there is a symmetry, then there exists a conservation law in the system. This is also known as Noether's theorem. (This only applies to continuous transformations.)

Of course, the big question is how do we define a symmetry? A transformation where the equations of motion are invariant? A geometric symmetry? An element of $SO(n)$?

Consider the generalized coordinates $q_i(t) \to Q_i(t, \lambda)$. We define it such that when $\lambda = 0, Q_i(t, 0) = q_i(t)$. So λ is simply some parameter. In our examples, a translation in $\hat{x}$ gives:

$$x(t) \to X(t, \lambda) = x(t) + \lambda \tag{53}$$

$$y(t) \to Y(t, \lambda) = y(t) \tag{54}$$

$$z(t) \to Z(t, \lambda) = z(t) \tag{55}$$

There is a subtlety in time translation (suppose a system that does not care about when it is being observed):

$$q_i(t) \to Q_i(t, \lambda) = q_i(t + \lambda) \tag{56}$$

This is because t is not a generalized coordinate. We are observing it at some different time.

Definition: If the equations of motion for Q_i's are the same as for the q_i's, then we say the system has a symmetry (i.e. You can replace Q_i with q_i and would be completely the same).

Recall that the Lagrangian of the system is not unique, and we are allowed to add a constant to the action and it would still yield the same equations of motion. This implies that we can add a total time derivative to our Lagrangian, allowing our equations of motion to remain invariant.

A transformation is a symmetry if and only if:

$$\frac{\partial \mathcal{L}(Q_i(t, \lambda), \dot{Q}_i(t, \lambda), t)}{\partial \lambda} = \frac{\mathrm{d}F}{\mathrm{d}t} \tag{57}$$

For some $F(Q_i, \dot{Q}_i, t)$. This implies that it leaves the EL equations unchanged.

Proof: Let $\mathcal{L} = \mathcal{L}(Q_i(t,\lambda), \dot{Q}_i(t,\lambda), t)$. Then looking at an infinitesimal transformation,

$$\frac{\partial \mathcal{L}}{\partial \lambda}|_{\lambda=0} = \sum_i \left[\frac{\partial \mathcal{L}}{\partial Q_i}\frac{\partial Q_i}{\partial \lambda} + \frac{\partial \mathcal{L}}{\partial \dot{Q}_i}\frac{\partial \dot{Q}_i}{\partial \lambda} \right]_{\lambda=0} \tag{58}$$

$$= \sum_i \left[\frac{\partial \mathcal{L}}{\partial q_i}\frac{\partial Q_i}{\partial \lambda} + \frac{\partial \mathcal{L}}{\partial \dot{q}_i}\frac{\partial \dot{Q}_i}{\partial \lambda} \right]_{\lambda=0} \tag{59}$$

Since $\frac{\partial \mathcal{L}}{\partial q_i} = \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial \mathcal{L}}{\partial \dot{q}_i}\right)$ by the EL equation, and $\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial Q_i}{\partial \lambda}\right) = \frac{\partial \dot{Q}_i}{\partial \lambda}$ and by the product rule this reduces to:

$$\frac{\partial \mathcal{L}}{\partial \lambda}|_{\lambda=0} = \sum_i \frac{\mathrm{d}}{\mathrm{d}t}\left[\frac{\partial \mathcal{L}}{\partial \dot{q}_i}\frac{\partial Q_i}{\partial \lambda}\right] = \frac{\mathrm{d}F}{\mathrm{d}t} \tag{60}$$

$$\therefore \frac{\mathrm{d}}{\mathrm{d}t}\left[\sum_i p_i \frac{\partial Q_i}{\partial \lambda}|_{\lambda=0} - F\right] = 0 \tag{61}$$

And we have shown that we are done, as there exists a conservation law in the system.

Examples: Spatial translations of closed systems. Consider a system of N particles that only interact with one another. The Lagrangian of a closed system is that there are no external forces.

So all coordinates are shifted by the same amount (under Cartesian coordinates). What is the transformation? Let $\vec{r_i}(t)$ indicate the position of the ith particle:

$$\vec{r_i}(t) \to \vec{r_i}(t) + \lambda \hat{n} = \vec{R_i}(t,\lambda) \tag{62}$$

If the Lagrangian is invariant, we can set $F = 0$ and it would be fine. Then, by Noether's theorem:

$$\sum_a p_a \frac{\partial Q_a}{\partial \lambda}|_{\lambda=0} \text{ is conserved} \tag{63}$$

Since $\frac{\partial \vec{R_i}}{\partial \lambda} = \vec{n}$, then:

$$\sum_{particles} \frac{\partial \mathcal{L}}{\partial \dot{\vec{r_i}}} \cdot \hat{n_i} = \sum_i m_i \vec{v} \cdot \hat{n} \tag{64}$$

$$\therefore \vec{P} = \sum_i \vec{p_i} = \sum m_i \vec{v_i} \tag{65}$$

Where $\vec{P}$ is the total amount of momentum of the system. This means that $\vec{P} \cdot \hat{n}$ is conserved. Since it is true for any $\hat{n}$, we can translate the whole system by $\hat{n}$ and $\vec{P}$ is invariant! (The *individual* momentum may not be conserved, but the *total* moment is conserved here).

7. Lecture 7 - Time and Rotational Invariance

Noether's theorem tells us something about the system without actually solving the equations of motion.

In particularly, time translation $q_i(t) \to q_i(t+\lambda) = Q_i(t,\lambda)$ means observing the system at some later time for the same generalized coordinate. What type of conservation exist in this system? The Lagrangian can depend on t implicitly (via $q(t), \dot{q}(t)$) and explicitly. We will focus more on the former case rather than the latter.

Let there be some $\mathcal{L} = \mathcal{L}(q_i, \dot{q}_i)$ where there is no explicit t dependence, then we have:

$$\frac{\mathrm{d}\mathcal{L}}{\mathrm{d}\lambda} = \frac{\mathrm{d}\mathcal{L}(Q_i(t,\lambda), \dot{Q}_i(t,\lambda))}{\mathrm{d}\lambda} = \frac{\mathrm{d}\mathcal{L}(q_i(t+\lambda), \dot{q}_i(t+\lambda))}{\mathrm{d}\lambda} \tag{66}$$

$$= \frac{\mathrm{d}\mathcal{L}(q_i(t+\lambda), \dot{q}_i(t+\lambda))}{\mathrm{d}t} = \frac{\mathrm{d}\mathcal{L}}{\mathrm{d}t} \Rightarrow F = \mathcal{L} \tag{67}$$

This implies that we have a conservation by definition of equating the first two lines. So anytime there is no explicit time dependence, we have a conservation. The quantity conserved is the generalized momentum.

$$\mathcal{H} = \sum_i \dot{q}_i p_i - \mathcal{L} \tag{68}$$

Where $p_i = \frac{\partial \mathcal{L}}{\partial q_i}$. The variable $\mathcal{H}$ is more commonly known as the Hamiltonian.

Claim: If the kinetic energy term in the Lagrangian involves purely quadratic terms, then the Hamiltonian represents the total energy of the system.

Proof: It uses Euler's theorem on homogeneous functions. A homogeneous function is defined as for some f:

$$f(\alpha x_1, ..., \alpha x_n) = \alpha^n f(x_1, ..., x_n) \tag{69}$$

For some n. If a function satisfies the above condition, then:

$$\sum_i \frac{\partial f}{x_i} = n f(x) \tag{70}$$

Let us now proceed with the proof. If T is quadratic in the $\dot{q}_i$'s,

$$\sum_i \dot{q}_i \frac{\partial T}{\partial \dot{q}_i} = 2T \tag{71}$$

Substituting into the Hamiltonian,

$$\mathcal{H} = \sum_i \dot{q}_i \frac{\partial T}{\partial \dot{q}_i} - \mathcal{L} = 2T - \mathcal{L} = 2T - (T - V) = T + V \tag{72}$$

We can recognize the right hand side of above to be the *total energy* of the system, something we recognize from QM. If the Lagrangian has explicitly time dependence, it typically represents that the total energy of the system is not conserved.

Now consider a rotation invariant system. This implies the conservation of angular momentum. Consider a rotation about the $\hat{n}$ axis, and an infinitesimally small rotation $\delta\vec{r}$ in our position $\vec{r}$. Note that $\delta\vec{r}$ has to be perpendicular to $\hat{n}$, and $\delta\vec{r}$ is perpendicular to $\vec{r}$.

We parametrize the rotation as $\delta\lambda$, where $\delta\lambda = \frac{|\delta\hat{r}|}{r\sin\theta}$.

$$\Rightarrow \delta\vec{r} = \hat{n} \times \vec{r}\delta\lambda \tag{73}$$

$$\Rightarrow \frac{d\vec{r}}{d\lambda} = \vec{n} \times \vec{r} \quad F = 0 \text{ (Lagrangian is unchanged)} \tag{74}$$

We plug this all into the Euler Lagrange equation, we will see that we have the $\hat{n}$ component of angular momentum is conserved. (Recall the spherical pendulum - n_x, n_y components of the angular momentum

is not conserved as it is not rotations about the x nor y axis)

$$\frac{\mathrm{d}}{\mathrm{d}t}\sum_i \vec{q_i}\cdot\frac{d\vec{r_i}}{\lambda}|_{\lambda=0} = \frac{\mathrm{d}}{\mathrm{d}t}\left[\sum_i \vec{p_i}(\hat{n}\times\vec{r_i})\right] = 0 \tag{75}$$

$$\Rightarrow \sum_i \vec{p_i}(\hat{n}\times\vec{r_i}) = \hat{n}\cdot\sum_i \vec{r_i}\times\vec{p_i} = \hat{n}\cdot\vec{M} \tag{76}$$

Where $\vec{M}$ is the total angular momentum of the system.

8. Lecture 8 - Mechanical Similarity and Integrals of Motion

Previously, we have shown that invariance under spatial translations, it implies that there is conserved linear momentum. For time translation invariance, we have the conserved Hamiltonian and under rotational translation invariance, there is conserved angular momentum.

We conclude that every closed system in $\mathbb{R}^3$ has at least 7 conserved quantities: $H(1), \vec{P}(3), \vec{M}(3)$. A simple pendulum is not a closed system, as the point of pivot holding the string to the ceiling does not enter the Lagrangian (there is an external force keeping it stuck to the roof). For a 2 dimensional pendulum, only the Hamiltonian is conserved - it is not spatially invariant, not invariant under rotations in θ. In 3 dimensions, it is invaiant under rotations about the z axis, so M_z is conserved. An example of a closed system would be two masses attached by a rod.

Comments to be made are that Noether's theorem gives us the idea of time invariant quantities. In quantum mechanics, the conservation of charge implies that you can multiply the wavefunction by a total phase. Another conservation law is the "Laplace-Runge-Lenz vector", defined as:

$$\vec{A} = \vec{p}\times\vec{M} - mk\hat{r} \tag{77}$$

This is only conserved in a Coulomb's potential. This corresponds to a symmetry of rotations on the surface of a 4-dimensional sphere, related to Kepler's problem. Conservation laws make the problem much more simpler.

If the system is not closed, some components of $\vec{p}$ and $\vec{M}$ may be conserved, and others are not. In a 3D pendulum, M_z is conserved but the other components are not. Similarly, if we have a falling ball the potential is $v = mgy$, we have that p_y is not conserved but the other components are.

Let us talk about the conservation of $\vec{p}$: Consider two reference frames, K and K' with the latter moving at velocity $\vec{v}$. Suppose a ball is thrown in K' with velocity $\vec{v_i}'$, then in K we observe the ball to move with velocity $\vec{v_i} = \vec{v_i}' + \vec{v}$.

The total momentum of the closed system is $\sum_i m_i\vec{v_i} = \sum_i m_i(\vec{v_i}) + \vec{v}\sum_i m_i$. That means:

$$\vec{p} = \vec{p'} + \vec{v}\sum_i m_i \tag{78}$$

This means there is actually a special choice of frame of reference we should choose when talking about rigid body motion. We can always choose K' such that $\vec{p'} = 0$ because of the conservation of momentum. The question is: How fast does that frame move?

If $\vec{P}'$ is zero, then the frame K' is moving at:

$$\vec{v} = \frac{\vec{P}}{\sum_i m_i} = \frac{\sum_i m_i \vec{v_i}}{\sum m_i} \tag{79}$$

$$= \frac{d}{dt}\left(\frac{\sum_i m_i \vec{r_i}}{\sum m_i}\right) \tag{80}$$

$$= \frac{d}{dt}\vec{R}_{center} \tag{81}$$

This is the center of mass of an object. So if we choose a frame such that the total momentum is zero, the center of mass is the origin of that frame. Therefore, the center of mass of a closed system movies with constant velocity. This will have important consequences, particularly about motion of rotating bodies.

Another important consequence: The total momentum and angular momentum is simply the sum of the momentum of each individual particle. However, the total energy of the system does not add linearly. The angular momentum of a system also depends on the choice of origin of the system (similar to dipole moment, the total angular momentum is only conserved under translations if the momentum of the object is zero). Consider a translation $\vec{r} \to \vec{r} + \vec{a}$, then:

$$\vec{M_i'} = \vec{r_i} \times \vec{p_i} \to (\vec{r_i} + \vec{a}) \times \vec{p_i} \tag{82}$$

$$= \vec{M_i} + \vec{a} \times \vec{p_i} \tag{83}$$

What is mechanical similarlity? In Noether's theorem, we have stated that a change in the coordinate system or adding a total time derivative gives us a conserved system if $\mathcal{L} \to \mathcal{L}$ or $\mathcal{L} \to \mathcal{L} + \frac{df}{dt}$. If $\mathcal{L} \to \alpha\mathcal{L}$ also leaves the equations of motion unchanged.

Suppose we have a potential that is homogeneous of degree k. In other words, $u(\alpha r_1, ..., \alpha r_n) = \alpha^k u(r_1, ..., r_n)$. For example, $u \propto \frac{1}{r}, u(\alpha r) = \alpha^{-1} u(r) \Rightarrow k = -1$. These are interesting because if we have a solution, we can scale our coordinates by a certain factor and get another solution.

For instance, we have $r_a \to \alpha r_a$ but we will also have to speed up or slow down our system at the same time, $t \to \beta t$. First question: What happens to our velocity? It will scale by $v_a = \frac{dr_a}{dt} \to \frac{\alpha}{\beta} v_a$.

This also means that our kinetic energy term T will scale by $\left(\frac{\alpha}{\beta}\right)^2$. On the other hand, the potential term U scales by α^k. If

$$\left(\frac{\alpha}{\beta}\right)^2 = \alpha^k \tag{84}$$

Then the Lagrangian $\mathcal{L} \to \alpha^k \mathcal{L}$ will be unchanged, as we are scaling everything by the same amount. If:

$$\Rightarrow \beta = \alpha^{1-\frac{k}{2}} \tag{85}$$

It leaves the system in variant. This means if $\vec{r_a}(t)$ is a solution to the equations of motion, then there is a whole family of solutions $\vec{r_a}'(t) = \alpha \vec{r_a}(\beta t)$.

For example, if particle moves a distance l in time t, there is another solution where it moves l' in t' where $\frac{t'}{t} = \left(\frac{l'}{l}\right)^{1-\frac{k}{2}}$.

9. Lecture 9 - Integrals of Motion

Let us assume a free particle in 3D. This gives us $\mathcal{L} = \frac{1}{2}m\vec{v}^2$. There will be 3 conserved quantities (technically 7) - linear momentum $\vec{p}(3)$, total energy $E(1)$ and angular momentum $\vec{M}(3)$. The initial conditions of the system are defined by only 6 quantities (the initial conditions are the positions and momentum), yet we have a total of 7 conserved quantities. This is because the energy is defined by $\vec{p^2}/2m$, and implies that one of the quantities is not independent.

In this case, for a single particle (only for a single particle are energy and momentum related) $E = \frac{\vec{p}^2}{2m}, \vec{p} \cdot \vec{M} = 0$ so we only have 5 independent conserved quantities.

Claim: A system with d degrees of freedom, it has a maximum number of conserved quantities of $2d - 1$. This is known as integrals of motion of conserved quantities.

If a system has d degrees of freedom, there are $2d$ quantities to be specified when giving initial conditions. However, if it is conserved w.r.t time, then we can reduce this by 1.

Consider a system with one degrees of freedom. We have $q, \dot{q}$. The conservation law here for m is;

$$f(q, \dot{q}) = 0 \quad [e.x. \frac{1}{2}m\dot{r}^2 + V(r) - E = 0] \tag{86}$$

This forms a curve in $(q, \dot{q})$ space (known as the phase space). The curve in phase space is a conservation law, where the particle is constrained to move along it. From this curve, we can determine the particle's trajectory without ever solving $\vec{F} = m\vec{a}$ as long as E is some constant. If we have another conserved quantity, we then have another curve in the phase space. We conclude that for a particle with 1 degree of freedom, it has at most 1 conserved quantity.

For a system with d degrees of freedom, the phase space will be $2d$ dimensional. The curve is specified by $2d - 1$ equations. Therefore, there is a max of $2d - 1$ independent integrals of motion.

a) Now consider 1 dimensional motion: $\mathcal{L} = \frac{1}{2}m\dot{x}^2 - U(x)$. From the Euler-Lagrange equation:

$$m\ddot{x} = -\frac{\mathrm{d}U}{\mathrm{d}x} \tag{87}$$

b) $\mathcal{H} = \frac{1}{2}m\dot{x}^2 + U(x)$ conserved. An example of this is the simple harmonic motion, where:

$$\dot{x}^2 + \frac{k}{m} = \frac{2E}{m} \tag{88}$$

Which is an elliptic curve in the phase space. We can solve for the motion without the Euler Lagrange equations, where we have:

$$\dot{x} = \frac{\mathrm{d}x}{\mathrm{d}t} = \sqrt{\frac{2(E - U(x))}{m}} \tag{89}$$

$$\Rightarrow \mathrm{d}t = \sqrt{\frac{m}{2}}\frac{\mathrm{d}x}{E - U(x)} \tag{90}$$

$$\Rightarrow \sqrt{\frac{m}{2}}\int \frac{\mathrm{d}x}{\sqrt{E - U(x)}} + C \tag{91}$$

This is known as “reduced to quadratures". The integrals of motion make things nicer. If we have d independent integrals of motion, we call the system “integrable". If we have $> d$, we call the system “super integrable". If we have $2d-1$ independent integrals of motion, then we call it “maximally integrable".

Let us revisit the $d=1$ example. We will find that for a system with conserved energy:

$$T = \sqrt{\frac{m}{2}} \int_{x_1(t)}^{x_2(t)} \frac{\mathrm{d}x}{\sqrt{E-U(x)}} \tag{92}$$

Nomenclature: x_1, x_2 will be referred to as turning points.

10. Lecture 10 - Central Forces and Potentials

Last time we discussed about integrable systems. For any one dimensional system, for a system with conserved energy:

$$\frac{1}{2}m\dot{x}^2 + U(x) = E \Rightarrow t = \sqrt{\frac{m}{2}} \int \frac{\mathrm{d}x}{\sqrt{E-U(x)}} \tag{93}$$

Which is referred to as “reduced to quadratures".

We have the problem of two particles in a central potential $U(\vec{r_1}, \vec{r_2}) = U(|\vec{r_1} - \vec{r_2}|)$, where U only depends on the distance between the two particles, not the direction. This gives us the Lagrangian:

$$\mathcal{L} = \frac{1}{2}m_1\vec{v_1}^2 + \frac{1}{2}m_2\vec{v_2}^2 - U(|\vec{r_1} - \vec{r_2}|) \tag{94}$$

There are 6 degrees of freedom of a closed system. All closed systems have conserved quantities (time translation, spatial translation, total energy). This implies that the center of mass moves with a constant velocity, and define it to be origin, which is the frame of reference we will be working in.

$$\vec{R}_{cen} = \frac{m_1\vec{r_1} + m_2\vec{r_2}}{m_1+m_2} = 0 \tag{95}$$

Define $\vec{r} = \vec{r_1} - \vec{r_2} \Rightarrow m_1\vec{r_1} + m_2\vec{r_2} = m1\vec{r} + (m_1+m_2)\vec{r_2} = 0$. This implies that:

$$\Rightarrow \vec{r_2} = \frac{-m_1}{m_1+m_2}\vec{r} \text{ and } \vec{r_1} = \frac{m_2}{m_1+m_2}\vec{r} \tag{96}$$

Where we reduce the position coordinates to in terms of $\vec{r}$. Note that since the center of mass stays the same, both objects are essentially orbiting around the origin. If we plug this back into the Lagrangian, we get:

$$\mathcal{L} = \frac{1}{2}\dot{\vec{r}}^2\left(\frac{m_1m_2}{m_1+m_2}\right) - U(r) \tag{97}$$

Where $r = |\vec{r}|$. Now if we examine the system closely, there seems to be only 3 degrees of freedoms for a particle which has an effective mass (or formally, the reduced mass) of $\mu = \frac{m_1m_2}{m_1+m_2}$ in an external potential $U(r)$.

If we can solve the problem of $\mathcal{L} = \frac{1}{2}\mu\dot{\vec{r}}^2 - U(r)$, then we have solved the original problem. Note that there are at most 5 degrees of freedom, and we know that all components of the angular momentum is conserved. However, linear momentum is not conserved as the potential depends on the distance. The

system is obviously time translation invariant, so the the energy (which is the Hamiltonian in this case) is conserved.

Since we have $4(>3)$ conserved quantities, this system is integrable, so we can solve using conservation laws. Since we have $\vec{M} = \vec{r} \times \vec{p}$ conserved, the particles are constrained to move in a single plane, as $\vec{r}, \vec{p}$ are perpendicular to $\vec{M}$. Now, we have a 2 dimensional problem.

Choose $\vec{M} = M_z \hat{z}$, so motion is constrained to the xy plane. Let us use spherical coordinates, with 2 degrees of freedom:

$$\mathcal{L} = \frac{1}{2}\mu(\dot{r}^2 + r^2\dot{\varphi}^2) - U(r) \tag{98}$$

Where M_z and E are conserved. Can we go down to 1 degree of freedom? We know that $M_z = (\vec{r} \times \vec{p})_z = \mu r^2 \dot{\varphi}$, so $\dot{\varphi} = \frac{M_z}{\mu r^2}$, where M_z is constant. Therefore, $\dot{\varphi} \propto \frac{1}{r^2}$. This allows us to reduce everything to 1 degree of freedom.

$$H = \frac{1}{2}\mu(\dot{r}^2 + r^2\dot{\varphi}^2) + U(r) = E \tag{99}$$

$$= \frac{1}{2}\mu\dot{r}^2 + \frac{M_z^2}{2\mu r^2} + U(r) \tag{100}$$

We can define an effective potential: Including any terms not carrying a $\dot{r}^2$.

$$U_{eff}(r) = U(r) + \frac{M_z^2}{2\mu r^2} \tag{101}$$

$$\Rightarrow H = \frac{1}{2}\mu\dot{r}^2 + U_{eff}(r) \tag{102}$$

So now we can say that the problem has been reduced to quadratures, as it is something much more recognizable from the last lecture.

Suppose $U(r)$ $\frac{-1}{r^\alpha}$ where $\alpha < 2$.

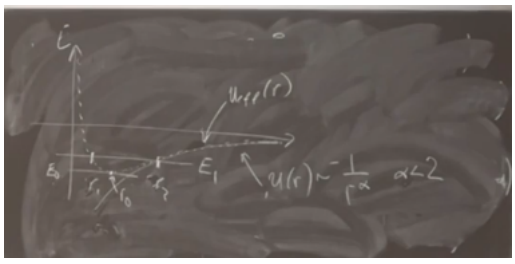


Figure 1: Complementary picture: At E_0, the particle moves in a circular orbit around the origin. At E_1, the particle would move in a spiral. There is no reason for orbits in a central potential to be closed.

We can find the radius of our orbit by finding the points of extrema of the effective potential.

11. Lecture 11 - Central Forces and Potentials (Continued)

Last lecture, we showed that in certain central potentials problems, it is equivalent to a one dimensional problem, allowing us to reduce to quadratures. When we think about orbits, where one mass is much

larger than the other, the reduced mass is essentially the mass of the smaller object. Also, $M_z = mr^2\dot{\varphi}$ is constant, so we find that the Hamiltonian is only dependent.

We can figure out the time required to do any orbit using the formula:

$$\frac{\mathrm{d}r}{\mathrm{d}t} = \sqrt{\frac{2}{m}(E - U_{eff}(r)} \tag{103}$$

$$\Rightarrow t = \int \frac{1}{\sqrt{\frac{2}{m}(E - U_{eff}(r)}} \mathrm{d}r \tag{104}$$

Note that the orbits are generally not closed. To get the angular motion, we need φ. We know that φ is a function of time, but we can write it as a function of r. So we want $\varphi(r)$, not $\varphi(t)$.

$$\frac{\mathrm{d}r}{\mathrm{d}t} = \frac{\mathrm{d}r}{\mathrm{d}\varphi}\frac{\mathrm{d}\varphi}{\mathrm{d}t} = \frac{M_z}{mr^2}\frac{\mathrm{d}r}{\mathrm{d}\varphi} \tag{105}$$

$$\Rightarrow \varphi(r) = \int \frac{M_z}{r^2}\frac{\mathrm{d}r}{\sqrt{2m(E - U_{eff}(r)}} \tag{106}$$

In particular, the change in φ in any one orbit we can solve for as an integral.

$$\Delta\varphi = 2\int_{r_1}^{r_2} \frac{M_z}{r^2}\frac{1}{\sqrt{2m(E - U_{eff}(r)}} \tag{107}$$

So $\Delta\varphi$ is not necessarily 2π.

Comments

We have the quantity $\frac{M_z^2}{2mr^2}$ is “really" just kinetic energy, in the scheme where we treat U_{eff} as potential energy. However, it actually does matter in the Lagrangian formalism $\mathcal{L} = T - V$. This can be done in the Hamiltonian, since we can't substitute $\dot{\varphi} = \frac{M_z}{mr^2}$ into $\mathcal{L}$ to get the equations of motion.

$$\Rightarrow \mathcal{L} = \frac{1}{2}m\dot{r}^2 + \frac{1}{2}\frac{M_z^2}{mr^2} - U(r) \tag{108}$$

Where the second term has a wrong sign!

Note that the term $\frac{M_z}{2mr^2}$ is a repulsive force. If we have a potential $U\ \frac{-1}{r^{2+n}}$ where $n > 0$, we find that the effective potential would be an attractive force.

A third comment to be made, suppose we have the sun (origin) and some mass m at distance $\vec{r}$ of our mass. Across some infinitesimally small time, it sweeps across some small angle $d\varphi$, meaning the mass moves by $rd\varphi$ during that time. This means the area it sweeps out during that time makes a triangle of area $dA = \frac{1}{2}r^2 d\varphi$.

$$\mathrm{d}A = \frac{1}{2}r^2\mathrm{d}\varphi \tag{109}$$

$$= \frac{1}{2}r^2\dot{\varphi}\mathrm{d}t \tag{110}$$

However, we know that $M_z = mr^2\varphi$ is constant for a central potential. This implies that:

$$\frac{\mathrm{d}A}{\mathrm{d}t} = \frac{1}{2}r^2\dot{\varphi} = \frac{M_z}{2m} \tag{111}$$

Which is the rate of change of area is the same. This is known as Kepler's 3rd law of planetary motion, which is actually just conservation of angular momentum in disguise.

Suppose we have a system with no potentials at all, and imagine the sun and earth. What would the effective potential be? We have $U = 0$, but we would have $U_{eff}(r) = \frac{M_z^2}{2mr^2}$ - but this is exactly a straight line, just that it is hard to see in the new coordinate system.

We have a special case of potentials: $\frac{1}{r}$ (gravity, Coulomb). Suppose we have the potential $U(r) = \frac{-k}{r}$. Then we get:

$$U_{eff}(r) = \frac{-k}{r} + \frac{M_z}{2mr^2} \tag{112}$$

Which gives us minima of $r = \frac{M^2}{mk}$. If we have that the enegy is less than 0, but greater than the minima, we will achieve an elliptical orbit. At precisely the minima, it travels at a circular orbit. If the energy was 0, then we have an unbounded orbit, like a comet that never returns.

12. Lecture 12 - Kepler's Laws

Given a central potential, we can define an effective potential,

$$U_{eff}(r) = \frac{M_z^2}{2mr^2} + U(r) \tag{113}$$

$$\Rightarrow \varphi(r) = \int \frac{M_z}{r^2} \frac{\mathrm{d}r}{\sqrt{2m(E - U_{eff}(r))}} \tag{114}$$

Which allows us to treat the problem as an effective 1-dimensional problem. Now suppose we have the potential $U(r) = \frac{-k}{r}$, and the Laplace-Runge-Lenz vector $\vec{A} = \vec{p} \times \vec{M} - mk\hat{r}$. We have shown in the problem set that this is a conserved quantity, so by Noether's theorem it must correspond to some form of symmetry. What kind? It is slightly non-trivial.

We have already shown $\frac{d\vec{A}}{dt} = 0$ on the problem set. For the full 3 dimensional problem, we have 3 degrees of freedom (think of a planet). It has maximum of 5 conserved quantities. For a central potential, we know that angular momentum and the total energy constitutes of this, and also the $\vec{A}$ vector, giving us 5 conserved quantities.

What is this vector $\vec{A}$? Consider the point where the orbit is the closest to the sun. What direction does $\vec{A}$ point in? It would be pointing radially outwards from the sun. Since it is a conserved quantity, it would always be pointing at the same direction regardless of the point of orbit.

Comments

Firstly, in the potential $\frac{1}{r}$, the orbit always has to be closed, since it would otherwise imply that $\vec{A}$ is not conserved as it would be pointing in a different direction for orbits that are not closed.

Secondly, suppose that the orbit is now in some other position $\vec{r}$, where the angle between $\vec{A}$ and $\vec{r}$ is φ. Then:

$$\langle \vec{A}, \vec{r} \rangle = Ar\cos\varphi = \vec{r} \cdot (\vec{p} \times \vec{M}) - mkr \tag{115}$$

$$= \vec{M} \cdot (\vec{r} \times \vec{p}) - mkr = M^2 - mkr \tag{116}$$

This is an algebraic equation. This implies that $M^2 = r(A\cos\varphi + mk)$, or rather:

$$\frac{1}{r} = \frac{mk}{M^2}(1 + e\cos\varphi) \tag{117}$$

Where $e = \frac{A}{mk} = \sqrt{1 + \frac{2EM^2}{mk^2}}$. The above equation gives us the shape of the orbit. This is precisely the equation of a conic section. The quantity "e" here is knkown as the eccentricity. If $e = 0$, then we have a circular orbit (as r is constant). If $e \leq 0 \leq 1$, then we have elliptical orbits. At $e = 1$, we have a parabolic orbit as $\frac{1}{r} \to 0$ at $\varphi = \pi$. When the energy $E > 0$, $r \to \infty$ for $\phi_0 < \pi$.

We have essentially solved Kepler's problem in a few lines.

(Midterm 1 covers everything before this).

Now we will move onto rigid body motion. When we throw an American football, two things happen: The football rotates around the axis of rotation, and the axis of rotation itself rotates around the football. It is actually a complicated type of motion.

We now consider the simplest type of rotational motion: A long pole, with two masses on it spinning at angular velocity Ω perpendicular to it. We have that M_z is conserved. This forms the equation:

$$M_z = I_z\Omega \tag{118}$$

Where I_z is a quantity known as the moment of inertia. Furthermore, $|\vec{r} \times \vec{p}| = rp = rmv_\theta = rmr\Omega$ since they are perpendicular. So the magnitude of angular momentum is simply $mr^2\Omega$.

$$\therefore M_z = 2mr^2\Omega = I_z\Omega \tag{119}$$

$$I_z = 2m\dot{r}^2 \tag{120}$$

(Factor of 2 since we have two masses on the pole). Let us not rotate about 1 symmetrical axis, and instead imagine some really weird axis of rotation. How do we analyze this sytem and figure out what the moment of inertia is?

Consider the rod making some angle θ with respect to the x-axis, rotating about the x-axis at angular velocity Ω. Let us define $\vec{\Omega}$, where it has the same magnitude as we know, but we define the direction using the right hand rule (with respect to direction of axis of rotation).

What is the angular momentum $\vec{M}$ of the system? It should be $\vec{r} \times \vec{p}$ as we know it. If we compute it, we find that $\vec{M}$ and Ω are not parallel, where the angular momentum is not conserved. This means there is a torque on the system - it doesn't like to spin like this. We **cannot** write it as $\vec{M} = I\vec{\Omega}$, where the moment of inertia will be a second rank tensor. They are related to the eigenvalues of a 3×3 matrix, and something known as the principal axis.

13. Lecture 13 - Dynamics of Extended Objects I

We have essentially been describing the motion of objects as if they were point masses, and determining their equations of motions. However, we failed to consider that these objects occupy some spatial volumes.

Objects can rotate around different axis, and it is possible that the axis of angular rotation is not parallel to the axis of angular momentum, so we cannot use simple equations such as $\vec{M} = I\vec{\Omega}$.

We want to talk about rigid bodies - we define rigid bodies to be masses which keeps their shape. **Definition:** An arrangement of masses which do not move relative to each other (keeps its shape).

Consider the density of some object $\rho(\vec{r})$, where the total mass is simply the volume the mass occupies multiplied by the distribution of density. $M = \int \rho(\vec{r})\mathrm{d}^3\vec{r}$.

Now let us consider the Lagrangian: How many degrees of freedom does a rigid body have? If it was a point object, it would have 3 degrees of freedom. The answer is 6. We first need to know the position of a single point (3 degrees of freedom). Then, we need to know how it spins with regards to the angles θ, φ, ψ (The two angles are the ones you already know, but ψ is the angle around the axis from the origin to the point on the mass).

When an object is in motion, we can always break down its motion into translational motion and rotational motion of the object. We will see this for a very convenient choice of a point on the object.

Consider 2 reference frames: The laboratory frame (inertial), and a frame attached to object (non-inertial in general). Consider the $(\hat{X}, \hat{Y}, \hat{Z})$ coordinate system of the laboratory frame, and the (x, y, z) for the coordiante system for the latter frame.

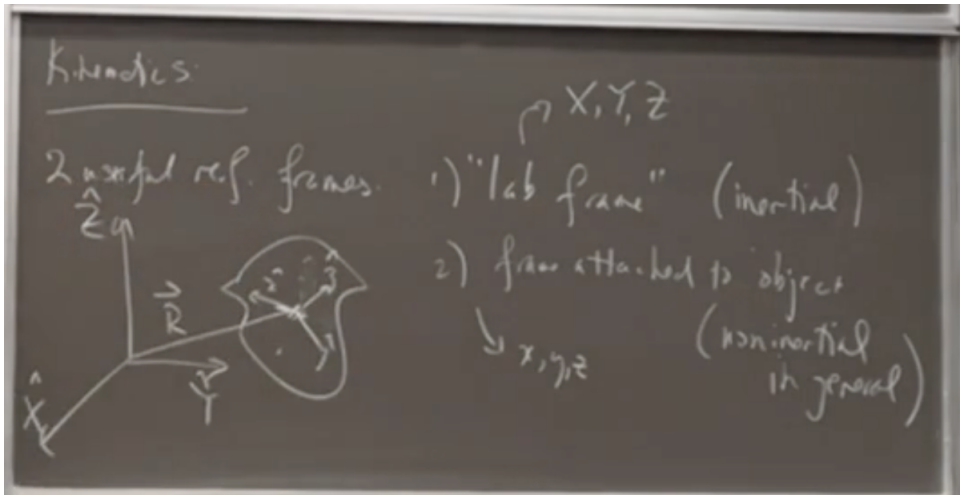


Figure 2: Reference picture, demonstrating the coordinate system.

Suppose we have some point on the mass, R and we fix some origin on our mass O. We call the vector from the origin (of the lab frame) to P to be $\vec{\rho}$, and the origin (of the lab frame) to the origin on the mass to be $\vec{R}$, and we define $\vec{r} = \vec{P} - \vec{R}$.

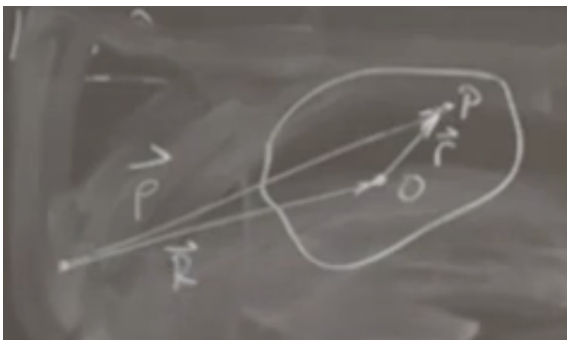


Figure 3: Imagine some mass in the laboratory frame - how do we describe its dynamics?

How does $\mathrm{d}t$ relate to $\mathrm{d}\rho$? We know that $|\vec{r}|$ has to be fixed by definition, since it is a rigid body. The movement of $\vec{r}$ has no translational motion, so it is simply a rotation. We can break up the motion into two pieces:

1. Displacement $\mathrm{d}\vec{R}$ of O
2. Rotation $\mathrm{d}\varphi$ about some axis $\vec{n}$ through O. As a notational note, we write $\mathrm{d}\vec{\varphi} = \mathrm{d}\varphi\hat{n}$.

By infinitesimals and some basic geometry, we have:

$$\mathrm{d}\vec{\rho} = \mathrm{d}\vec{R} + \mathrm{d}\vec{r} \tag{121}$$
$$= \mathrm{d}\vec{R} + \mathrm{d}\varphi \times \vec{r} \tag{122}$$

The former term refers to the motion as a whole, and the latter to be the rotational motion. Dividing both sides by $\frac{1}{\mathrm{d}t}$,

$$\Rightarrow \underbrace{\frac{\mathrm{d}\vec{\rho}}{\mathrm{d}t}}_{\text{Velocity of } P} = \underbrace{\frac{\mathrm{d}\vec{R}}{\mathrm{d}t}}_{\text{Velocity of } O} + \underbrace{\frac{\mathrm{d}\vec{\varphi}}{\mathrm{d}t} \times \vec{r}}_{\text{Velocity of } P \text{ wrt } O} \tag{123}$$

We can write the above as $\vec{v} = \vec{V} + \vec{\Omega} \times \vec{r}$ where $\vec{\Omega} = \frac{\mathrm{d}\vec{\varphi}}{\mathrm{d}t}$, which tells us the velocity of any point on our object. How should we choose O? There are infinitely many points on a space occupying object, and two different people might define different origins on the mass: O and O'. Let us denote the vector from O to O' to be $\vec{a}$.

$$\vec{R'} = \vec{R} + \vec{a} \tag{124}$$
$$\vec{\rho} = \vec{R} + \vec{r} = \vec{R} + \vec{r'} = \vec{R} + \vec{a} + \vec{r'} \tag{125}$$

Immediately we get that:

$$\vec{v} = \vec{V} + \vec{\Omega} \times (\vec{r'} + \vec{a}) \tag{126}$$
$$= \vec{V} + \vec{\Omega} \times \vec{a} + \vec{\Omega} \times \vec{r'} \tag{127}$$

but we have also have $\vec{v} = \vec{V'} + \vec{\Omega'} \times \vec{r'}$, and both are equal for any O we choose on the mass (V' is the velocity of O'). This gives us the relation

$$\vec{V'} = \vec{V} + \vec{\Omega} \times \vec{a} \tag{128}$$

$$\Rightarrow \vec{\Omega'} = \vec{\Omega} \tag{129}$$

In other words, regardless of the choice of origin on the mass, the angular velocity is independent of it. It is a property of the object, not a property of the coordinate system we choose.

Notes:

- The instantaneous angular velocity of an object is independent of the origin.
- We can choose a point O' for which $V' = 0$, which we refer to the instantaneous axis of rotation.

Example: Consider a circular wheel rolling on a table. What is the motion of point P?

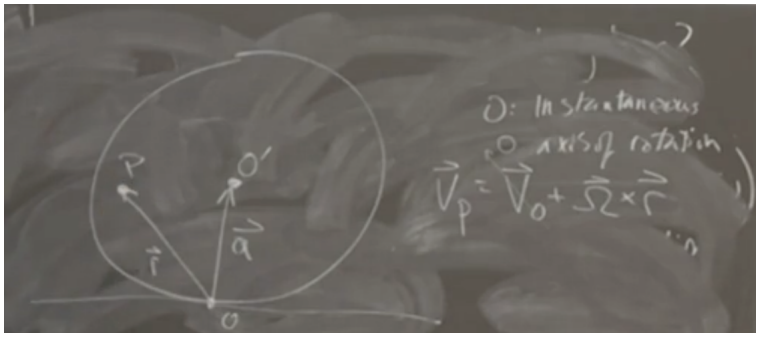


Figure 4: O is an instantaneous axis of rotations

We can write that $\vec{v}_p = \vec{v}_o + \vec{\Omega} \times \vec{r} = \vec{v}'_o + \vec{\Omega} \times \vec{r'}$, where $\vec{r'}$ is the vector pointing from O' to P. What we find that is that we can find the motion in two different ways, and we get the same velocity.

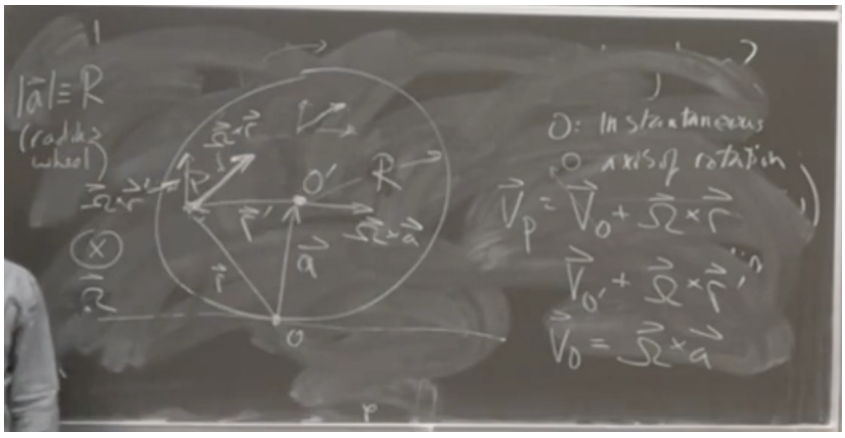


Figure 5: The kinematics very much depends on what we choose the origin to be.

14. Lecture 14 - Dynamics of Extended Objects II

Last time, we discussed about the fact that $\vec{p} = \vec{v_O} + \vec{\Omega} \times \vec{r}$, where $\vec{\Omega}$ is independent of the choice of O. The first term refers to the velocity of the origin, and the velocity of the rotation around the origin.

Let us write down the Lagrangian, in a scenario where $V = 0$. We have $\mathcal{L} = T$ - what is the kinetic energy of a rotating object? Take the object to be made of a large number of point masses m_a, then

$$T = \frac{1}{2} \sum_a m_a v_a^2 \tag{130}$$

Where $\vec{v_a} = \vec{v} + \vec{\Omega} \times \vec{r_a}$

$$\Rightarrow \mathcal{L} = \frac{1}{2} \sum_a m_a (\vec{v}^2 + 2\vec{v} \cdot (\vec{\Omega} \times \vec{r_a}) + (\Omega \times \vec{r_a})^2) \tag{131}$$

We can choose a nice choice of O to make the cross term zero by choosing O to be the instantaneous axis of rotation. Another way is to choose O to be the centre of mass, and note that (triple product ideentity)

$$\vec{v} \cdot (\vec{\Omega} \times \vec{r_a}) = \vec{r_a} \cdot (\vec{v} \times \vec{\Omega}) \tag{132}$$

Giving us $(\sum m_a \vec{r_a}) \cdot (\vec{v} \times \vec{\Omega})$. The terms in the parenthesis would vanish, since at the centre of mass, the weighted sum of all the $\vec{r_a}$ would cancel out and vanish, and hence the cross term would go to zero(?).

For the 3rd term of our Lagrangian, there is a quadruple vector identity that states:

$$(\vec{A} \times \vec{B}) \cdot (\vec{C} \times \vec{D}) = (\vec{A} \cdot \vec{C})(\vec{B} \cdot \vec{D}) - (\vec{A} \cdot \vec{D})(\vec{B} \cdot \vec{C}) \tag{133}$$

So we can simplify the Lagrangian to:

$$\mathcal{L} = \underbrace{\frac{1}{2} M v^2}_{\text{KE of translation}} + \underbrace{\frac{1}{2} \sum_a m_a (\Omega^2 r_a^2 - (\vec{\Omega} \cdot \vec{r_a})^2)}_{\text{KE of rotation}} \tag{134}$$

Where $M = \sum m_a$, where the above is equivalent to the kinetic energy of a rotating object. Consider the coordinate system $\hat{x_i}$ for $i = 1, 2, 3$ (forget about relativistic mechanics).

We have $r_a^2 = \sum_i r_{ai}^2$. This gives:

$$(\vec{\Omega} \cdot \vec{r_{ai}})^2 = (\sum_i \Omega_i r_{ai})(\sum_j \Omega_j r_{aj}) \tag{135}$$

$$= \sum_{ij} \Omega_i \Omega_j r_{ai} r_{aj} \tag{136}$$

Note that $\Omega^2 = \sum_i \Omega_i \Omega_i = \sum_{ij} \Omega_i \Omega_j \delta_{ij}$. Plugging this back into the rotational kinetic energy, we get:

$$T_{rot} = \frac{1}{2} \sum_{ij} \Omega_i \Omega_j \left[\sum_a m_a (\delta_{ij} \vec{r_a}^2 - r_{ai} r_{aj}) \right] \tag{137}$$

Let us pause for a minute: Why is it a good way to write this thing? This looks much more convoluted than what we had originally. Consider the term in the square brackets: It is something that only depends on the distribution of the mass of the object, not the velocity of the object. The terms in the square bracket

is known as the “moment of inertia", which will be denoted as I_{ij}. It is a second rank tensor.

$$I_{ij} = \sum_a m_a(\delta_{ij} r_a^2 - r_{ai} r_{aj}) \tag{138}$$

This forms a symmetrical matrix - it has 6 free numbers. Let us compare I_{ij} with a vector - vectors typically have 3 components (where $i = 1, 2, 3$), but this tensor has 9 components $(i, j = 1, 2, 3)$, where i denotes the row and j denotes the column.

Note that:

$$\sum_{ij} \Omega_i \Omega_j I_{ij} \iff \begin{pmatrix} \Omega_1 & \Omega_2 & \Omega_3 \end{pmatrix} \begin{pmatrix} I_{11} & I_{12} & I_{13} \\ I_{21} & I_{22} & I_{23} \\ I_{31} & I_{32} & I_{33} \end{pmatrix} \begin{pmatrix} \Omega_1 \\ \Omega_2 \\ \Omega_3 \end{pmatrix} \tag{139}$$

The expression on the left is index notation, which we will primarily be using moving forward. In its full glory, the moment of inertia tensor:

$$I = \sum_a \begin{pmatrix} m_a(y_a^2 + z_a^2) & -m_a x_a y_a & -m_a x_a z_a \\ -m_a x_a y_a & m_a(x_a^2 + z_a^2) & -m_a y_a z_a \\ -m_a x_a z_a & -m_a y_a z_a & m_a(x_a^2 + y_a^2) \end{pmatrix} \tag{140}$$

Note that $I_{ij} = I_{ji}$. Suppose we have a continuous mass distribution with density $\rho(\vec{r})$, then the components of the moment of inertia tensor becomes:

$$I_{ij} = \int_{\text{Object}} \rho(\vec{r})(\vec{r}^2 \delta_{ij} - r_i r_j) \mathrm{d}^3 \vec{r} \tag{141}$$

However, I_{ij} actually depends on the choice of axis $\hat{x_i}$, so in general $I_{ij} \neq I'_{i'j'}$ (try flipping the $\hat{x}$ and $\hat{y}$ axis around and see that the matrix above is trivially not equal). Are there some choices of axis nicer than others for the moment of inertia tensor? These are called the principle axis, and in these axis, we find that I becomes diagonal. This is a problem equivalent to finding the eigenvalues and eigenvectors of I. What do these eigenvectors represent?

15. Lecture 15 - Principal Axis I

The principal axis are a set of coordinate vectors such that the moment of inertia tensor is diagonal. For rotations about a principal axes, $\vec{M}$ and $\vec{\Omega}$ are parallel.

If we work in a basis where the motion of inertia tensor is not diagonal, then it is complicated. For a real symmetrical matrix in $\mathbb{R}^3$, we know it has three orthogonal eigenvectors, so it is possible to find such a diagonal basis. For any object (regardless of the geometric symmetry), there always exists at least one set of orthogonal axes for which $I_I ij$ is diagonal. In that basis, we only need I_1, I_2, I_3 to write out the moment of inertia tensor. This basis is called the principal axes, and the eigenvalues are the “principal moments of inertia".

Proof: In linear algebra, we know that we can diagonalize a symmetric matrix by an orthogonal transformation.

$$I^{diag} = OIO^T \tag{142}$$

Where $OO^T = O^T O = \mathbb{I}$, and O is an orthogonal matrix. It is evident that $I_{1,2,3}$ are the eigenvalues of I.

In principle, given the moment of inertia matrix, we can go ahead and find the eigenvalues and the eigenvectors to find the principal axes and the principal moments of inertia. Now, let us think about this physically.

Consider the angular momentum of a rigid body (e.g. American football). By definition of angular momentum, the total angular momentum of it is given by:

$$\vec{M} = \sum_a (\vec{r}_a \times \vec{p}_a) = \sum_a m_a (\vec{r}_a \times \vec{v}_a) \tag{143}$$

O is either the center of mass, or the instantaneous axis of rotation, so we have $\vec{v}_a = \vec{\Omega} \times \vec{r}_a$. Recall the triple cross product identity $\vec{a} \times (\vec{b} \times \vec{c}) = \vec{b}(\vec{a} \cdot \vec{c}) - \vec{c}(\vec{a} \cdot \vec{b})$.

$$\Rightarrow \vec{r}_a \times (\vec{\Omega} \times \vec{r}_a) = \vec{\Omega} r_a^2 - \vec{r}_a \vec{\Omega} \cdot \vec{r} \tag{144}$$

Substituting this into the equation for total angular momentum,

$$M_i = \sum_j \Omega_j \underbrace{\left[\sum_a m_a (\vec{r}_a^2 \delta_{ij} - r_{ai} r_{aj}) \right]}_{I_{ij}} \tag{145}$$

So to get the vector $\vec{M}$, we multiply the vector Ω by the matrix I.

$$\begin{pmatrix} M_1 \\ M_2 \\ M_3 \end{pmatrix} = \begin{pmatrix} & I_{ij} & \end{pmatrix} \begin{pmatrix} \Omega_1 \\ \Omega_2 \\ \Omega_3 \end{pmatrix} \tag{146}$$

In general, $\vec{M}$ and $\vec{\Omega}$ are not parallel. If we spin it about a fixed Ω, then $\vec{M}$ will rotate about $\vec{\Omega}$. We will need an external force for this to happen, as we will need an external torque to change the angular momentum.

The angular velocity $\vec{\Omega}$ is the vector that the object is spinning around. However, that vector itself also rotates in certain situations, so the axis of rotation changes in order to the keep the angular momentum fixed.

Supposing we took an American football and spun it about one of its principal axes. What happens? It turns out that the angular velocity vector will be parallel to the angular momentum vector. Why is that?

Well, I will be diagonal, so $M_i = I_i \Omega_i$ for $i = 1, 2, 3$. This implies that the object will not wobble as the components of angular momentum will always be some constant multiple of the components of angular velocity. So if $\Omega_2 = \Omega_3 = 0$, then $\vec{M}$ and $\vec{\Omega}$ are parallel, so $\vec{M}$ stays constant as object rotates and spins freely.

Let us suppose we took two masses, a distance of $2a$ apart each with mass m and spun it about some axis that is not the principal axis. What are the components of the moment of inertia tensor? Imagine that it lies in the x, y plane, and we spun it about its midpoint making an angle θ with respect to the x-axis.

$$I_{11} = m(a^2 - a^2\cos^2\theta) + m(a^2 - a^2\cos^2\theta) \tag{147}$$
$$= 2ma^2\sin^2\theta \tag{148}$$
$$I_{12} = -2ma^2\sin\theta\cos\theta \tag{149}$$
$$I_{22} = 2ma^2\cos^2\theta \tag{150}$$
$$I_{13} = I_{23} = 0 \tag{151}$$
$$I_{33} = 2ma^2 \tag{152}$$

This forms the moment of inertia tensor,

$$I = 2ma^2\begin{pmatrix} \sin^2\theta & -\sin\theta\cos\theta & 0 \\ -\sin\theta\cos\theta & \cos^2\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{153}$$

Which is clearly not diagonal. What are the principal axis? Obviously we can make it diagonal when $\theta = 0$. So the principal axis of this system are precisely the x, y, z axes.

General observations::

1. Consider any object with a plane of symmetry. Then we know that there is always a principal axis perpendicular to it, and the other two are in the plane of symmetry.
2. Consider any object with an axis of symmetry. Then, the axis of symmetry is a principal axis, and the other two are perpendicular to the axis (and there are an infinite number of choice for them).
3. For a spherical tops, any set of axes are the principal axes.

16. Lecture 16 - Principal Axis II

The principle moments of inertia tells us the mass distribution of the object.

For example, a football is perfectly spherical so we can spin it about any axis and we won't see any wobbling, as any axis is a principal axis for it.

We classify objects like the following:

1. **Spherical Top** - All of the moment of inertia are the same, so $I_1 = I_2 = I_3$. This implies the moment of inertia is simply proportional to the identity matrix.
2. **Symmetric Top** - Two of the moment of inertias are the same, so $I_1 = I_2 \neq I_3$ (e.g. American football).
3. **Asymmetric Top** - All three moments of inertias are not equal.

For any object,

$$I_1 + I_2 = \sum_a (y_a^2 + z_a^2 + x_a^2 + z_a^2) m_a \tag{154}$$

$$= \sum_a m_a (x_a^2 + y_a^2 + 2z_a^2) \tag{155}$$

$$\geq I_3 \tag{156}$$

So the moment of inertia cannot exceed the sum of the other two. Another comment to be made is that I_{ij} depends on the origin O. Fortunately, we can relate the two matrices. Suppose we know I_{ij} with respect to the center of mass. Suppose O' is O displaced by some vector $\vec{b}$. Then,

$$I'_{ij} = \sum m_a (\delta_{ij} {r'_a}^2 - r'_{ai} r'_{aj}) \tag{157}$$

$$= \sum_a m_a (\delta_{ij} [{r_a}^2 - 2\vec{b} \cdot \vec{r}_a + \vec{b}^2] - r_{ai} r_{aj} + r_{ai} b_j + b_i r_{aj} - b_i b_j) \tag{158}$$

However, but the initial origin being the center of mass implies $\sum_a m_a \vec{r}_a = 0$. Then,

$$\Rightarrow I'_{ij} = \sum_a m_a (\delta_{ij} r_a^2 - r_{ai} r_{aj}) + \sum_a m_a (b^2 \delta_{ij} - b_i b_j) \tag{159}$$

The first term is simply the moment of inertia about the center of mass, and the second term is some term involving the total mass. This is known as the parallel axis theorem.

$$I'_{ij} = I_{ij} + \mu (b^2 \delta_{ij} - b_i b_j) \tag{160}$$

Where μ is the total mass of the object. Note that a cube is technically a spherical top, since its principle moments of inertia are all equal. Let the side length of the cube be equal to l, and the density $\rho = \frac{M}{l^3}$. Then its moment of inertia components (choosing the origin to be the center of mass):

$$I_{11} = \int_{-\frac{l}{2}}^{\frac{l}{2}} \int_{-\frac{l}{2}}^{\frac{l}{2}} \int_{-\frac{l}{2}}^{\frac{l}{2}} \rho (y^2 + z^2) \mathrm{d}x \mathrm{d}y \mathrm{d}z \tag{161}$$

$$= Ml^2/6 \tag{162}$$

By symmetry, we know that $I_{11} = I_{22} = I_{33}$, so we have the moment of intertia tensor:

$$I = \frac{Ml^2}{6} \mathbb{I} \tag{163}$$

What if we chose the origin to be the corner of the cube instead? This means we shifted it by the vector $\vec{b} = -\frac{l}{2}\hat{x} - \frac{l}{2}\hat{y} - \frac{l}{2}\hat{z}$.

Which implies that $b^2 = \frac{3l^2}{4}, b_i b_j = l^2/4$ for any i, j. By the parallel axes theorem,

$$I'_{ij} = Ml^2 \begin{pmatrix} \frac{2}{3} & -\frac{1}{4} & -\frac{1}{4} \\ -\frac{1}{4} & \frac{2}{3} & -\frac{1}{4} \\ -\frac{1}{4} & -\frac{1}{4} & \frac{2}{3} \end{pmatrix} \tag{164}$$

This means that the typical $\hat{x}, \hat{y}, \hat{z}$ axes that are principal at the center of mass are not principal along the corner of the cube, or as you say the angular momentum is not parallel to the axis of rotation.

Let us look at the free rotations of rigid bodies, working in the center of mass frame (so we only have rotational energy in the kinetic term.

$$\mathcal{L} = T = \frac{1}{2}\sum_{ij} \Omega_i I_{ij} \Omega_j \tag{165}$$

Where $\Omega_i = \frac{\mathrm{d}\theta_i}{\mathrm{d}t}$ (generalized velocities). Let us look at the Euler-Lagrange equations.

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial \mathcal{L}}{\partial \Omega_k}\right) = 0 \tag{166}$$

By the product rule,

$$\Rightarrow \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{1}{2}\sum_{ij}\left[\underbrace{\frac{\partial \Omega_i}{\partial \Omega_k}}_{\delta_{ik}} I_{ij}\Omega_j + \Omega_i I_{ij} \underbrace{\frac{\partial \Omega_j}{\partial \Omega_k}}_{\delta_{jk}}\right]\right) = 0 \tag{167}$$

If we re-label the indices, what we get as the equation of motion is:

$$\frac{\mathrm{d}}{\mathrm{d}t}\sum_j I_{kj}\Omega_j = \frac{\mathrm{d}}{\mathrm{d}t} M_k = 0 \tag{168}$$

So the Euler-Lagrange equation simply says the angular momentum is conserved for a free body, where no external forces are present. This is not a surprise from what we have already learnt.

For a spherical top, suppose we spin it about $\hat{x}_3$, then $M_{1,2} = 0$. We know that the angular velocity has to be constant, and the object will spin freely at Ω. What about a symmetrical top?

If we spun a symmetrical top (e.g. American football) along some axis not parallel to one of the principal axes. Let $\Omega_2 = 0$ for the sake of argument (and the other components of angular velocity are non-zero). So $\vec{M} = I_1\Omega_1\hat{x}_1 + I_3\Omega_3\hat{x}_3$. If $I_1 > I_3$, then $\vec{M}$ points closer to $\hat{x}_1$ than $\hat{x}_3$, so $\vec{\Omega}$ is not parallel to $\vec{M}$. Also, $\vec{\Omega}$ is not constant, but $\vec{M}$ is (since it is a free body). This means that $\vec{\Omega}$ will rotate about $\vec{M}$, and points on the object will rotate about $\vec{\Omega}$. This is known as torque-free precession.

17. Lecture 17 - Euler Equations

In the previous lecture, we discussed about the torque-free precession of a symmetric top. There exists a fixed angular momentum $\vec{M}$, but there is also $\vec{\Omega}$ and $\hat{x_3}$ that lies coplanar to it in plane S at all times. The angular momentum **has** to be fixed in a closed system.

We showed that $\vec{v_a} = \vec{\Omega} \times \vec{r_A}$. What we find that the axis of rotation, $\vec{\Omega}$ and $\hat{x_3}$ will also rotate around $\vec{M}$ (precession, this occurs if the axis of symmetry does not align with the angular momentum vector). What we conclude is that the plane S rotates uniformly about $\vec{M}$.

Be careful: $\hat{x_3}$ rotates around $\vec{\Omega}$ by definition, but $\vec{\Omega}$ also rotates around $\vec{M}$. What happens in the end with the conservation of momentum is the two orbits around $\vec{M}$.

How do we determine how the object precesses? There are two methods to approach this: A laboratory frame, and the fixed body frame. We will start with the second frame: A frame of reference where the principle axis are fixed.

Euler Equations

This is useful for the fixed body frame, and imagine with have some object with principal axis $(\hat{x_i})$. Under free rotation (no external forces), we have $\frac{\mathrm{d}\vec{M}}{\mathrm{d}t} = 0$ in the **lab frame** is fixed. However, the components of the angular momentum in the fixed body frame will NOT be fixed in the lab frame, since the principal axis rotate with respect to time.

$$\frac{\mathrm{d}\vec{M}}{\mathrm{d}t} \Rightarrow \sum_i \frac{\mathrm{d}}{\mathrm{d}t}\left[M_i(t)\hat{x}_i(t)\right] \tag{169}$$

$$= \sum_i \left[\frac{\mathrm{d}M_i}{\mathrm{d}t}\hat{x_i} + M_i \underbrace{\frac{\mathrm{d}\hat{x_i}}{\mathrm{d}t}}_{\vec{\Omega}\times\hat{x}_i}\right] = 0 \tag{170}$$

For a fixed vector $\vec{a}$ on an object, $\frac{d\vec{a}}{dt} = \Omega \times \vec{a}$. We immediately get:

$$\Rightarrow \frac{\mathrm{d}\vec{M}}{\mathrm{d}t} = \sum_i \left[\frac{\mathrm{d}M_i}{\mathrm{d}t}\hat{x}_i + M_i(\vec{\Omega}\times\hat{x}_i)\right] \tag{171}$$

Based on the formula for the cross product $(\vec{a}\times\vec{b})_1 = a_2b_3 - a_3b_2$, we have that:

$$\therefore \frac{\mathrm{d}M_1}{\mathrm{d}t} - M_2\Omega_3 + M_3\Omega_2 = 0 \tag{172}$$

Around the principal axis, we have $M_i = I_i\Omega_i$. This gives us the ultimate relation:

$$I_1\dot{\Omega_1} + \Omega_2\Omega_3(I_3 - I_2) = 0 \tag{173}$$

This tell sus how it changes in the rotating frame. We can do the exact same thing in the $\hat{x_2}$ and $\hat{x_3}$ components. This generates a system of differential equations:

$$I_2\dot{\Omega}_2 + \Omega_3\Omega_1(I_1 - I_3) = 0 \tag{174}$$

$$I_3\dot{\Omega}_3 + \Omega_1\Omega_2(I_2 - I_1) = 0 \tag{175}$$

Together, they form the Euler equations for free rotation (where we see the RHS equals to 0). This works for any rigid body - let us apply this to a symmetrical top, where $I_1 = I_2 \neq I_3$ (an American football). This immediately tells us that:

$$\dot{\Omega}_1 = \Omega_2\Omega_3(\frac{I_1 - I_3}{I_1}) \tag{176}$$

$$\dot{\Omega}_2 = -\Omega_1\Omega_3(\frac{I_1 - I_3}{I_1}) \tag{177}$$

$$\dot{\Omega}_3 = 0 \tag{178}$$

These are the equations of motion for the spinning top, and immediately tells us that Ω_3 is constant in this frame. Now take $\Omega_1, \Omega_2 \neq 0$,

$$\Rightarrow \dot{\Omega}_1 = -\omega\Omega_2 \tag{179}$$

$$\dot{\Omega}_2 = -\omega\Omega_1 \tag{180}$$

Where $\omega = \Omega_3(\frac{I_3 - I_1}{I_1})$ is constant in this case. The solution to this is simple harmonic motion:

$$\Rightarrow \ddot{\Omega}_1 = -\omega\dot{\Omega}_2 = -\omega^2\Omega_1 \tag{181}$$

$$\ddot{\Omega}_2 = -\omega^2\Omega_2 \tag{182}$$

So that explains why it wobbles around - Ω_1, Ω_2 oscillates with frequency ω: It's simple harmonic motion! If $\Omega_2 = 0$ at $t = 0$,

$$\Omega_1 = \Omega_0 \cos\omega t \tag{183}$$

$$\Omega_2 = \Omega_0 \sin\omega t \tag{184}$$

In the body fixed frame, $\vec{\Omega}$ precesses about the $\hat{x}_3$ axis with frequency ω.

Example: "Chandler Wobble"

- The Earth is technically a symmetrical top, since it bulges slightly at the equators due to its rotations. It has $a\hat{x}_3$, with $\frac{I_3 - I_1}{I_1} \approx \frac{1}{300}$. The Earth is also not spinning on its principle axis, where $\hat{x}_3$ is not aligned with $\vec{\Omega}$. What we find is that the Earth's axis of rotation precesses about $\hat{x}_3$

How fast does the Earth rotate? We can find this using how long a day is: $\Omega_3 = \frac{2\pi}{day}$, then the precessional frequency has $\omega \approx \frac{2\pi}{300}\text{day}^{-1}$, giving us a precessional period of ≈ 300 days (experimentally it is 427 days).

We claim the whole system precesses about $\vec{M}$ (in the laboratory frame). In the body fixed frame, Ω precesses about $\hat{x}_3$. We ask the question: In the lab frame, how fast does Ω precess around $\vec{M}$?

The plane S is rotating at $\omega\hat{x}_3$ in a frame which is rotating at $\vec{\Omega}$ in the lab frame. Therefore, the plane S is rotating at $\Omega + \omega\hat{x}_3$ in the laboratory frame. (We can just add angular velocities together). Let us work this out rigorously.

At the instant that S is on $x_1 - x_3$ plane,

$$\vec{\Omega} + \omega\hat{x}_3 = \Omega_1\hat{x}_1 + \Omega_3\hat{x}_3 + \Omega_3\left(\frac{I_3 - I_1}{I_1}\right)\hat{x}_3 \tag{185}$$

$$= \Omega_1\hat{x}_1 + \frac{I_3}{I_1}\Omega_3\hat{x}_3 \tag{186}$$

$$= \frac{1}{I_1}\underbrace{(I_1\Omega_1\hat{x}_1 + I_3\Omega_3\hat{x}_3)}_{\vec{M}} \tag{187}$$

$$= \frac{\vec{M}}{I_1} \tag{188}$$

Therefore in the lab frame, the whole thing processes at $\vec{\Omega}_{pr} = \frac{\vec{M}}{I}$.

18. Lecture 18 - Asymmetrical tops and Euler Angles

Last day we approached the equations of motions for a top (e.g. American football) in the rotating frame of the object in terms of the generalized velocities Ω_i, giving us the components of Ω_i along the principal axis. Now, let us examine what's going on in the laboratory frame. Euler angles are used to describe the kinematics in the lab frame.

Is rotation about a principal axis (or a small perturbation around it) stable? Let us suppose we spin a tennis racket along $\hat{x}_1$ axis, then:

$$\Omega_1 = \Omega \text{ constant} \tag{189}$$
$$\Omega_2 = 0 \tag{190}$$
$$\Omega_3 = 0 \tag{191}$$

Now suppose we have a slight perturbation (so $\eta_i << \Omega$):

$$\Omega_1 = \Omega + \eta_1(t) \tag{192}$$
$$\Omega_2 = \eta_2(t) \tag{193}$$
$$\Omega_3 = \eta_3(t) \tag{194}$$

When we put this back into the Euler equations, we get the system of equations,

$$I_1\dot{\eta_1} = \eta_2\eta_3(I_3 - I_2) = O(\eta^2) \tag{195}$$

So $\dot{\eta}_1 \approx= 0$. For the other two equations, we get:

$$I_2\dot{\eta_2} \approx \Omega\eta_3(I_3 - I_2) \tag{196}$$
$$I_3\dot{\eta_3} \approx \Omega\eta_2(I_1 - I_2) \tag{197}$$

Taking the time derivative on both $\dot{\eta}$ equations,

$$\Rightarrow I_2\ddot{\eta_2} = \frac{\Omega^2}{I_3}(I_3 - I_1)(I_1 - I_2)\eta_2 \tag{198}$$
$$= A\eta_2 \tag{199}$$

Where we define $A = \frac{\Omega^2(I_3-I_1)(I_1-I_2)}{I_3}$. We can recognize this to be a 2nd order ODE. Hence, if $A < 0$ then we get sinusoidal solutions for η, and for $A > 0$ this would be unstable as it would give exponential solutions of η, therefore unstable.

If $(I_3 - I_1)(I_1 - I_2) > 0$, then we have $I_3 > I_1$ and $I_1 > I_2$ **or** $I_3 < I_1$ and $I_1 < I_2$. This implies that:

$$I_2 < I_1 < I_3 \tag{200}$$
$$I_2 > I_1 > I_3 \tag{201}$$

This is known as the "intermediate axis theorem"/tennis racket theorem/Dzanibekov Effect, where rotations about the intermediate axis are unstable. For a symmetrical top, they are both stable.

The idea for Euler angles is that we do not want to talk about Ω_i's but instead of angular velocities that make sense in the lab frame. We have X, Y, Z (lab frame axis), and we define θ, φ as typical from spherical coordinates. Lastly, we define ψ as the angle around $\hat{x}_3$, the third principal axis. The problem is to relate $\Omega_1, \Omega_2, \Omega_3$ to $\dot{\theta}, \dot{\varphi}, \dot{\psi}$. The tennis racket has the principal axis $\hat{x}_1, \hat{x}_2, \hat{x}_3$.

Recall rotations in 2D: Suppose two coordinate systems in there (x, y) and (x', y') differ by angle θ. Then:

$$\hat{x}' = \cos\theta\hat{x} + \sin\theta\hat{y} \tag{202}$$

$$\hat{y}' = -\sin\theta\hat{x} + \cos\theta\hat{y} \tag{203}$$

We can rewrite this in matrix form,

$$\begin{pmatrix} \hat{x}' \\ \hat{y}' \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} \hat{x} \\ \hat{y} \end{pmatrix} \tag{204}$$

The rotation matrix is trivially an element in $SO(2)$, where we obtain the relation $R(\theta_2)R(\theta_1) = R(\theta_2 + \theta_1)$ for any θ. This is also true in $\mathbb{R}^3$ (though commutativity does not hold - this relation forms $SO(3)$).

Let $R_z(\theta)$ denote the rotations about the z-axis in $\mathbb{R}^3$. Since it leaves components in the z-axis unchanged, then its matrix is simply represented as:

$$R_Z(\theta) = \begin{pmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{205}$$

Likewise, for rotations about the X axis, it simply are planar rotations in the YZ plane. This generates the matrix:

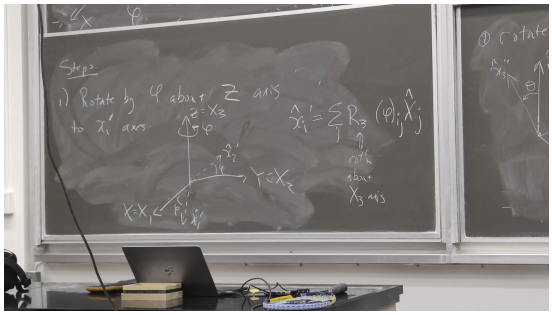
$$R_X(\theta) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & \sin\theta \\ 0 & -\sin\theta & \cos\theta \end{pmatrix} \tag{206}$$

The steps for the full process are as follow:

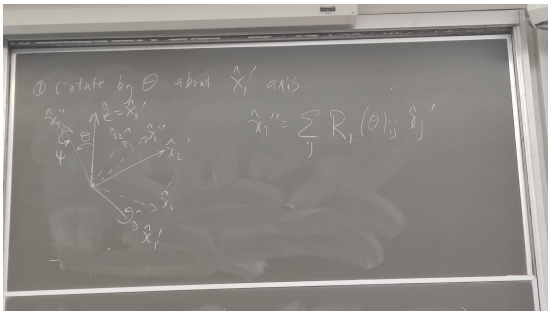
1. Rotate ψ about the Z-axis to x_i' axis. Using index notation we have the relation $\hat{x}_i' = R_3(\varphi)_{ij}\hat{x}_j$ between the new and old axis. φ is the angle corresponding between $\hat{x}_i'$ and $\hat{x}_i$.
2. Rotate by θ about the $\hat{X}_1'$ axis, and use the relation $\hat{x}_i'' = R_1(\theta_{ij})\hat{x}_j'$, where θ is the angle corresponding between $\hat{x}_3''$ and $\hat{x}_3$.
3. Finally, rotate by ψ about $\hat{x}_3''$ to get the body fixed axis. This gives us the transformation rule $\hat{x}_i = R_3(\psi)_{ij}\hat{x}_j''$

We claim that:

$$\begin{pmatrix} \hat{x}_1 \\ \hat{x}_2 \\ \hat{x}_3 \end{pmatrix} = R_3(\psi)R_1(\theta)R_3(\varphi) \begin{pmatrix} \hat{X} \\ \hat{Y} \\ \hat{Z} \end{pmatrix} \tag{207}$$



(a) Corresponding to step (1)



(b) Corresponding to steps (2) and (3)

What's left is simply a matrix multiplication calculation (slightly tedious). This yields that:

$$R(\theta,\varphi,\psi)=\begin{pmatrix}\cos\psi\cos\varphi-\cos\theta\sin\varphi\sin\psi & \sin\varphi\cos\psi+\cos\theta\sin\psi\cos\varphi & \sin\theta\sin\psi\\ -\cos\varphi\sin\psi-\cos\theta\cos\psi\sin\varphi & -\sin\psi\sin\varphi+\cos\theta\cos\psi\cos\varphi & \sin\theta\cos\psi\\ \sin\theta\sin\varphi & -\sin\theta\cos\varphi & \cos\theta\end{pmatrix} \tag{208}$$

19. Lecture 19 - Euler angles and Laboratory Frames

How do we relate the Euler angles to the principal axes of the object? We stated this last lesson.

1. ψ around $\hat{Z}$, where $(\hat{X},\hat{Y},\hat{Z})\to(\hat{x}_1',\hat{x}_2',\hat{x}_3')$
2. θ around $\hat{x}_1'$, where $(\hat{x}_1',\hat{x}_2',\hat{x}_3')\to(\hat{x}_1'',\hat{x}_2'',\hat{x}_3'')$
3. φ around $\hat{x}_3''$, where $(\hat{x}_1'',\hat{x}_2'',\hat{x}_3'')\to(\hat{x}_1,\hat{x}_2,\hat{x}_3)$

We want to write our angular velocities in terms of our Euler angles. We note that our Euler angles will change with respect to time.

$$\Rightarrow\vec{\Omega}=\dot{\varphi}\hat{Z}+\dot{\theta}\hat{x}_1'+\dot{\varphi}\hat{x}_3 \tag{209}$$

This is not a basically a nice form, since we are mashing different basis together. Let us write them out in terms of the principal axes ($\hat{x}_3$ is already one of them). How do we write $\hat{Z},\hat{x}_1'$ in terms of $\hat{x}_1,\hat{x}_2,\hat{x}_3$? We will need to write out the inverse of $R(\theta,\varphi,\psi)$ so we have the relation:

$$\begin{pmatrix}\hat{X}\\ \hat{Y}\\ \hat{Z}\end{pmatrix}=R^{-1}(\psi,\theta,\varphi)\begin{pmatrix}\hat{x}_1\\ \hat{x}_2\\ \hat{x}_3\end{pmatrix} \tag{210}$$

By symmetry, R^{-1} is actually quite trivial to obtain - we simply reverse the order of rotations (and by a negative angle). Then $R^{-1}=R(-\varphi,-\theta,-\psi)$. Form matrix multiplication, we find that:

$$\hat{Z}=\sin\theta\sin\psi\hat{x}_1+\sin\theta\cos\psi\hat{x}_2+\cos\theta\hat{x}_3 \tag{211}$$

$$\hat{X}_1'=\cos\psi\hat{x}_1-\sin\psi\hat{x}_2 \tag{212}$$

Where the second one comes from rotating $\hat{x}_1$ by $-\psi$ about the $\hat{x}_3$ axis to get $\hat{x}_1'$. This generates the system of equations for the angular velocity $\vec{\Omega} = \Omega_1\hat{x}_1 + \Omega_2\hat{x}_2 + \Omega_3\hat{x}_3$ (where $\hat{x}_i$ are the principal axes).

$$\Omega_1 = \dot{\varphi}\sin\theta\sin\psi + \dot{\theta}\cos\psi \tag{213}$$
$$\Omega_2 = \dot{\varphi}\cos\psi - \dot{\theta}\sin\psi \tag{214}$$
$$\Omega_3 = \dot{\psi} + \dot{\varphi}\cos\theta \tag{215}$$

If $\theta(t) = 0$, then $\Omega_1 = \Omega_2 = 0, \Omega_3 = \dot{\psi} + \dot{\varphi}$. This is simply a constant rotation about the $\hat{Z}$ axis.

If $\theta(t) = \frac{\pi}{2}$: The $\hat{x}_3$ axis is perpendicular to the $\hat{Z}$ axis, and it spins around it at $\dot{\psi}$ and $\hat{x}_3$ is precessing about $\hat{Z}$ at $\dot{\varphi}$.

For a symmetrical top (American football), let us define $\hat{Z}$ to be the direction of the angular velocity (which we know to be fixed). The motion around the $\hat{Z}$ axis is $\dot{\varphi}$, the motion around $\hat{x}_3$ is $\dot{\psi}$, and the angle $\hat{x}_3$ makes with respect to $\hat{Z}$ is θ.

$$M_1 = I_1\Omega_1 \tag{216}$$
$$M_2 = I_2\Omega_2 = I_1\omega_2 \tag{217}$$
$$M_3 = I_3\Omega_3 \tag{218}$$

We choose when $\psi = 0$ and look at that instant. This immediately implies that $\hat{x}_1' = \hat{x}_1$, and we have rotations about θ are rotations about $\hat{x}_1$ axis. For $\psi = 0$ instantaneously,

$$\Omega_1 = \dot{\theta} \tag{219}$$
$$\Omega_2 = \dot{\varphi}\sin\theta \tag{220}$$
$$\Omega_3 = \dot{\varphi}\cos\theta + \dot{\psi} \tag{221}$$

From here, we note that $\hat{x}_1 \perp \vec{M} = M_1 = 0 \Rightarrow \Omega_1 = 0 \Rightarrow \dot{\theta} = 0$. It has precession frequency $\Omega_{pr} = \dot{\varphi}$.

$$\dot{\varphi} = \frac{\Omega_2}{\sin\theta} = \frac{M_2}{I_2\sin\theta} \tag{222}$$
$$\text{since } \mathrm{M}_2 = \mathrm{M}\sin\theta, \quad = \frac{M\sin\theta}{I_2\sin\theta} \tag{223}$$
$$= \frac{M}{I_2} = \frac{M}{I_1} \tag{224}$$

What about $\dot{\psi}$? It is the angular frequency of $\hat{x}_{1,2}$ axes about $\hat{x}_3$ (we called this ω before).

$$\Omega_3 = \frac{M_3}{I_3} = \dot{\varphi}\cos\theta + \dot{\psi} = \frac{M\cos\theta}{I_1} + \dot{\psi} \tag{225}$$
$$\text{since } \mathrm{M}\cos\theta = \mathrm{M}_3, \quad \Rightarrow \dot{\psi} = \frac{M_3}{I_3} - \frac{M_3}{I_1} \tag{226}$$
$$= \Omega_3\left(\frac{I_3 - I_1}{I_1}\right) = \omega \tag{227}$$

20. Lecture 20 - Heavy Symmetrical Top with Gravity

For a gyroscope (heavy symmetrical top with gravity), we will be writing down everything we have learnt so far. The Euler angles are very convenient.

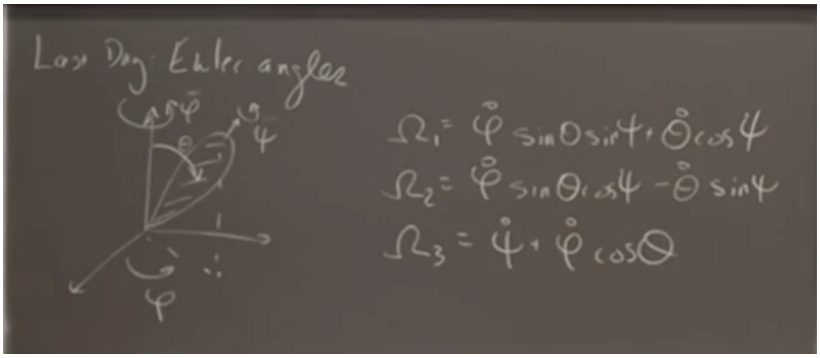


Figure 7: Euler angles of a rotating object

What is the Lagrangian of the system? This is a symmetrical top, so $I_1 = I_2$. Let us measure things about P, the origin. The height of the center of mass is then given by $l\cos\theta$

$$\mathcal{L} = \frac{1}{2}I_1(\Omega_1^2 + \Omega_2^2) + \frac{1}{2}I_3\Omega_3^2 - \mu g l\cos\theta \tag{228}$$

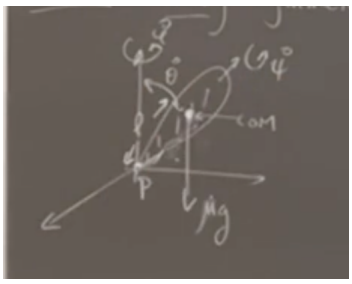


Figure 8: Reference picture of what's going on

Now we need to write Ω in terms of the Euler angles. We will have to shift our coordinates. Doing the algebra and plugging things in,

$$\Rightarrow \mathcal{L} = \frac{1}{2}I_1(\dot{\theta}^2 + \sin^2\theta\dot{\varphi}^2) + \frac{1}{2}I_3(\dot{\psi} + \cos\theta\dot{\varphi})^2 - \mu g l\cos\theta \tag{229}$$

We have something in terms of 3 Euler angles, so we can apply the Euler-Lagrange equations and obtain the equations of motion if we wish. We can analyze the system before we start doing this though, as it saves us a lot of work. Are there any conserved quantities?

If we look at the Lagrangian, it really only depends on velocities, and the only coordinate that shows up is θ. We recognize that ψ is a cyclic variable, which means its conjugate momentum ($\partial\mathcal{L}/\partial\dot{\psi}$)is conserved. The Hamiltonian is also conserved as there is no explicit time dependence. The system is invariant under rotations along the ψ axis, and also invariant under rotations under the φ axis.

- $\mathcal{L}$ is independent of ψ and φ, so $p_\psi = \frac{\partial \mathcal{L}}{\partial \dot{\psi}}$ and $p_\varphi = \frac{\partial \mathcal{L}}{\partial \dot{\varphi}}$ are conserved. Be careful - ψ is cyclic but $\dot{\psi}$ is not conserved.
- That is because what is actually conserved is the conjugate momentum $p_\psi = I_3(\dot{\psi} + \cos\theta\dot{\varphi}) = I_3\Omega_3$, as our axis are not necessarily orthogonal. So Ω_3 is conserved.
- The other quantity $p_\varphi = \frac{\partial \mathcal{L}}{\partial \dot{\varphi}} = I_1 \sin^2\theta\dot{\varphi} + I_3\cos\theta(\dot{\psi} + \dot{\varphi}\cos\theta) = I_1\sin^2\theta\dot{\varphi} + I_3\Omega_3\cos\theta$ is conserved.
- Invariant under time translations, as the Hamiltonian has no time dependence and since it is quadratic in the generalized velocities, the Hamiltonian is the total energy is conserved for this system. $E = T + V = \frac{1}{2}I_1(\dot{\theta}^2 + \dot{\varphi}^2\sin^2\theta) + \frac{1}{2}I_3\Omega_3^2 + \mu g l\cos\theta$.
- Since we have 3 conserved quantities, and three degrees of freedom, the system is integrable, meaning we can solve this system using conservation laws (without diving into EL).

Let us solve this problem using conservation laws. Let us solve this under a special case: Is there a solution to these equations of motion where θ is constant? It appears so (under a physical demo done in the class). To show this, skipping a couple of lines given that total energy is conserved

$$\frac{\mathrm{d}E}{\mathrm{d}t} = 0 \text{ (or } \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial \mathcal{L}}{\partial \dot{\theta}}\right) = \frac{\partial \mathcal{L}}{\partial \theta}) \tag{230}$$

$$\Rightarrow I_1\ddot{\theta} = (\mu g l + I\dot{\varphi}^2\cos\theta - I_3\Omega_3\dot{\varphi})\sin\theta \tag{231}$$

Then it must be that $\dot{\theta} = 0$, immediately implying that the terms in the parenthesis above must equal to zero if a solution exists where θ is constant. Let's see:

$$\dot{\varphi} = \frac{I_3\Omega_3}{2I_1\cos\theta}\left(1 \pm \sqrt{1 - \frac{4\mu I_1 g l\cos\theta}{I_3^2\Omega_3^2}}\right) \tag{232}$$

This is the precession frequency. We know Ω_3 is conserved by our previous steps. There are two solutions to this equation referred to as the fast and slow precession. The requirement that there exists a solution is whatever under the square root is > 0.

- If $\cos\theta > 0$ ($\theta < \frac{\pi}{2}$), the top is pointing up and there is a minimum frequency $\Omega_3 = \frac{2}{I_3}\sqrt{I_1\mu g l\cos\theta}$ below which there is no solution (as the top falls).

For a fast top, $I_3\Omega_3 >> \sqrt{I_1\mu g l}$. Then:

$$\Rightarrow \dot{\varphi}_{slow} \approx \frac{\mu g l}{I_3\Omega_3} \tag{233}$$

$$\dot{\varphi}_{fast} \approx \frac{I_3\Omega_3}{I_1\cos\theta} \tag{234}$$

The one we "always" physically see is the slow solution. Why is that? In the classroom, we would have to set up $\dot{\varphi}$ to go very fast in order for that to be true, which is not achievable by standard means (unless we are working with some certain equipment).

To solve this in the general case, we will add a nutation (so $\dot{\theta} \neq 0$). Then:

$$p_\psi = I_3 \Omega_3 \tag{235}$$

$$p_\varphi = I_1 \sin^2\theta \dot{\varphi} + I_3 \varphi_3 \cos\theta \tag{236}$$

$$E = \frac{1}{2} I_1 (\dot{\theta}^2 + \dot{\varphi}^2 \sin^2\theta) + \frac{1}{2} I_3 \Omega_3^2 + \mu g l \cos\theta \tag{237}$$

These are our constraints, as they are conserved quantities. We define $a = \frac{p_\psi}{I_1} = \frac{I_3 \Omega_3}{I_1}$ and $b = \frac{p_\psi}{I_1}$, where we can still see that both of them are conserved quantities. Plugging them back into our previous expressions,

$$\Rightarrow \dot{\varphi} = \frac{b - a\cos\theta}{\sin^2\theta} \tag{238}$$

$$\Rightarrow \dot{\psi} = \frac{I_1 a}{I_3} - \frac{(b - a\cos\theta)\cos\theta}{\sin^2\theta} \tag{239}$$

It turns out that $\dot{\psi}$ also changes as it bobs up and down along θ. If we substitute this into E, we get:

$$E' = E - \frac{1}{2} I_3 \Omega_3^2 \tag{240}$$

Where the new energy E' is also conserved. Then we can write that:

$$\Rightarrow E' = \frac{1}{2} I_1 \dot{\theta}^2 + V_{eff}(\theta) \tag{241}$$

What is this effective potential? We work this out to be:

$$V_{eff}(\theta) = I_1 \frac{(b - a\cos\theta)^2}{2\sin^2\theta} + \mu g l \cos\theta \tag{242}$$

The effective potential is a one dimensional problem dependent on θ. There is a value of θ for which $b = a\cos\theta$, so $\dot{\varphi}$ can oscillate back and forth too given a small initial nudge. The plot of the effective potential looks like:

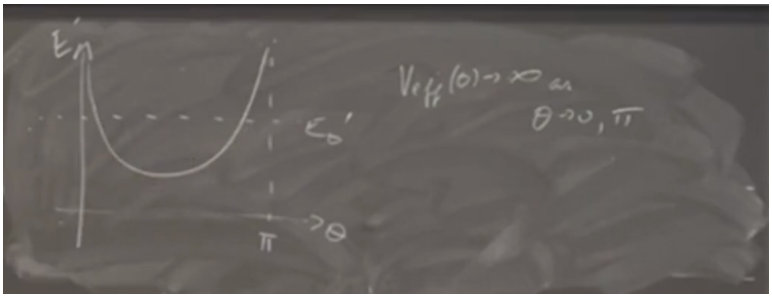


Figure 9: Given energy E_0', it will oscillate like it does for a 1 dimensional problem, bobbing up and down. There is a minimum, where θ doesn't change.

When $\dot{\theta}$ doesn't change, we get pure precession. This is the minimum point of the effective potential as seen on the plot. Reducing the problem to quadratures,

$$t = \int \frac{\mathrm{d}\theta}{\sqrt{2(E' - V_{eff}(\theta))/I_1}} \tag{243}$$

Unfortunately, this is an elliptic integral, so we instead have to understand it qualitatively.

21. Lecture 21 - Hamiltonian Formalism

These are known as the canonical equations. For the Lagrangian formalism, we know that for $i = 1, ..., n$ $\mathcal{L}(q_i, \dot{q}_i, t)$ generates the equations of motion using the Euler-Lagrange equations.

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial \mathcal{L}}{\partial \dot{q}_i}\right) = \frac{\partial \mathcal{L}}{\partial q_i} \tag{244}$$

We can re-write kinematics in a more elegant way that is totally equivalent. We have seen that $\mathcal{L} \iff \vec{F} = m\vec{a}$, and EL can make a set of n 2nd order PDE's. In order to determine unique solutions to this set of differential equations, $2n$ initial conditions must be given.

We have defined a quantity known as generalized momenta: $p_i = \frac{\partial \mathcal{L}}{\partial \dot{q}_i}$ where $i = 1, ..., n$, then $\dot{p}_i = \frac{\partial \mathcal{L}}{\partial q_i}$. The idea is that the in the Lagrangian formalism, both $q, \dot{q}_i$ must be given. In the Hamiltonian formalism, we want to describe the whole system in terms of q, p (not $\dot{q}$). Is there a function $\mathcal{H}(q_i, p_i, t)$ which plays a role similar to $\mathcal{L}$ but gives the equations of motion in terms of q_i, p_i? The answer is yes (and there are two).

Consider $\mathcal{H}(q_i, p_i, t) = \sum_i p_i \dot{q}_i - \mathcal{L}(q_i, \dot{q}_i, t)$. We have already learnt about this equation previously when learning about conservation laws. Note that $\mathcal{H}$ is strictly in terms of (q_i, p_i, t), but $\mathcal{L}$ carries $\dot{q}_i$. This means we will have to invert the relation:

$$p_i = \frac{\partial \mathcal{L}}{\partial \dot{q}_i} \longrightarrow \text{ given } p_i(q_i, \dot{q}_i, t) \rightarrow \text{ invert this} \tag{245}$$

This gives $\dot{q}_i(q_i, p_i, t) \Rightarrow$ so this is only a function of q_i, p_i, t and not $\dot{q}_i$. How do we generate the equations of motion from the Hamiltonian $\mathcal{H}$? Consider a small variation:

$$\mathrm{d}\mathcal{H} = \sum_i \left(p_i \mathrm{d}\dot{q}_i + \dot{q}_i \mathrm{d}p_i - \frac{\partial \mathcal{L}}{\partial q_i}\mathrm{d}q_i - \frac{\partial \mathcal{L}}{\partial \dot{q}_i} d\dot{q}_i\right) - \frac{\partial \mathcal{L}}{\partial t}\mathrm{d}t \tag{246}$$

$$\text{since } \left(p_i = \frac{\partial \mathcal{L}}{\partial \dot{q}_i}\right), \ = \sum_i \left(\dot{q}_i \mathrm{d}p_i - \frac{\partial \mathcal{L}}{\partial q_i}\mathrm{d}q_i\right) - \frac{\partial \mathcal{L}}{\partial t}\mathrm{d}t \tag{247}$$

$$\text{since } \left(\dot{p}_i = \frac{\partial \mathcal{L}}{\partial q_i}\right), \ = \sum_i (\dot{q}_i \mathrm{d}p_i - \dot{p}_i \mathrm{d}q_i) - \frac{\partial \mathcal{L}}{\partial t}\mathrm{d}t \tag{248}$$

However, just by applying the chain rule

$$\mathrm{d}\mathcal{H} = \sum_i \left(\frac{\partial \mathcal{H}}{\partial q_i}\mathrm{d}q_i + \frac{\partial \mathcal{H}}{\partial p_i}\mathrm{d}p_i\right) + \frac{\partial \mathcal{H}}{\partial t}\mathrm{d}t \tag{249}$$

By reading the like terms, then:

$$\Rightarrow \dot{q}_i = \frac{\partial \mathcal{H}}{\partial p_i}, \ \dot{p}_i = -\frac{\partial \mathcal{H}}{\partial q_i} \text{ and } \frac{\partial \mathcal{H}}{\partial t} = -\frac{\partial \mathcal{L}}{\partial t} \tag{250}$$

These are known as Hamilton's equations. They are expressed in terms of q_i, p_i **not** $q_i, \dot{q}_i$. They are nicely symmetric in the p_i's and the q_i's.

For example, look at a particle in a potential. Then,

$$\mathcal{L} = \frac{1}{2}m\dot{\vec{r}}^2 - V(\vec{r}) \tag{251}$$

Obtain the p_i, q_i:

$$p_i = \frac{\partial \mathcal{L}}{\partial \dot{r}_i} = m\dot{r}_i \Rightarrow \dot{r}_i = \frac{p_i}{m} \tag{252}$$

Hence,

$$\mathcal{H} = \vec{p} \cdot \dot{\vec{r}} - \mathcal{L} \tag{253}$$

$$= \frac{1}{m}\vec{p}^2 - \frac{1}{2}m\left(\frac{\vec{p}^2}{m^2}\right) + V(\vec{r}) \tag{254}$$

$$= \frac{\vec{p}^2}{2m} + V(\vec{r}) \tag{255}$$

We can now entirely forget about the Lagrangian,

$$\dot{r}_i = \frac{\partial \mathcal{H}}{\partial p_i} = \frac{p_i}{m} \tag{256}$$

$$\dot{p}_i = -\frac{\partial \mathcal{H}}{\partial r_i} = -\frac{\partial V}{\partial r_i} \rightarrow \dot{\vec{p}} = -\vec{\nabla} V(\vec{r}) \tag{257}$$

Which are consistent with the Lagrangian formalism and Newton's second law.

Let us look at another example: The simple pendulum of length l with some angle with respect to the rest position θ. Then we know the Lagrangian is:

$$\mathcal{L} = \frac{1}{2}ml^2\dot{\theta}^2 + mgl\cos\theta \tag{258}$$

$$p_\theta = \frac{\partial \mathcal{L}}{\partial \dot{\theta}} = ml^2\dot{\theta} \tag{259}$$

$$\mathcal{H} = \frac{p_\theta^2}{2ml^2} - mgl\cos\theta \tag{260}$$

Thus, we obtain that:

$$\dot{p}_\theta = -\frac{\partial \mathcal{H}}{\partial \theta} = -mgl\sin\theta \tag{261}$$

$$\dot{\theta} = \frac{\partial \mathcal{H}}{\partial p_\theta} = \frac{p_\theta}{ml^2} \tag{262}$$

Hamilton's equations generates $2n$ 1st order partial differential equations, compared to EL's n 2nd order partial differential equations. This means both of them require $2n$ initial conditions. We have went from $\mathcal{L}(q_i, \dot{q}_i, t) \rightarrow \mathcal{H}(q_i, p_i, t)$ (This is mathematically known as a "Legendre transform", though we won't care so much about it in this class). Why do we care?

It is much more nicer to numerically solve a system of 1st order DE's than a system of 2nd order DE's (though it is not for this reason we care so much about it in this course, this is a theoretical class).

For theoretical reasons, there is something known as the phase space flow. Think of the coordinates $\{q_i, p_i\}$, and that $\dot{p}_i = -\frac{\partial \mathcal{H}}{\partial q}, \dot{q}_i = \frac{\partial \mathcal{H}}{\partial p_i}$. The phase space of a simple pendulum:

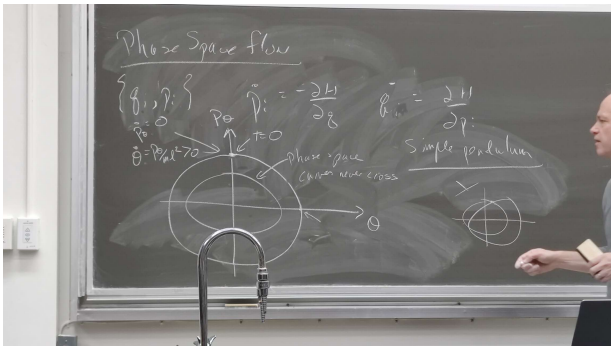


Figure 10: At $t = 0, \dot{p}_\theta = 0, \dot{\theta} = \frac{p_\theta}{ml^2} > 0$ by Hamilton's equations (we previously obtained them already). Choosing the initial condition where $\dot{p}_\theta = 0$ generates a closed trajectory in the phase space (the larger circle). It turns out for a simple pendulum, we will always obtain a closed curve.

Phase space curves can never intersect, as it would violate Picard's theorem (since it implies the same initial condition can generate two different trajectories).

It is convenient to think of solutions to Hamilton's equations as a unique trajectory of the phase space of $\{q_i, p_i\}$. Let us propose the question: As we move along a fixed trajectory in the phase space, how do the functions of q_i, p_i as the system evolves in time?

For the energy ($\frac{\partial \mathcal{L}}{\partial t} = 0$),

$$\frac{\mathrm{d}\mathcal{H}(q_i, p_i, t)}{\mathrm{d}t} = \sum_i \frac{\partial \mathcal{H}}{\partial q_i}\dot{q}_i + \frac{\partial \mathcal{H}}{\partial p_i} + \underbrace{\frac{\partial \mathcal{H}}{\partial t}}_{=0 \text{ because } \frac{\partial \mathcal{L}}{\partial t}=0} \tag{263}$$

$$= \sum_i -\dot{p}_i\dot{q}_i + \dot{q}_i\dot{p}_i = 0 \tag{264}$$

They are also cyclic,

$$\frac{\partial \mathcal{H}}{\partial q_i} = 0 \Rightarrow \dot{p}_i = 0 \tag{265}$$

Lastly, for $f(q_i, p_i, t)$:

$$\frac{\mathrm{d}f}{\mathrm{d}t} = \sum_i \left[\frac{\partial f}{\partial p_i}\dot{q}_i + \frac{\partial f}{\partial q_i}\dot{q}_i\right] + \frac{\partial f}{\partial t} \tag{266}$$

$$= \sum_i \underbrace{\left[-\frac{\partial f}{\partial p_i}\frac{\partial \mathcal{H}}{\partial q_i} + \frac{\partial f}{\partial q_i}\frac{\partial \mathcal{H}}{\partial p_i}\right]}_{\text{Poisson Bracket}} + \frac{\partial f}{\partial t} \tag{267}$$

$$= \{f, \mathcal{H}\} + \frac{\partial f}{\partial t} \tag{268}$$

We refer to the terms in the square brackets as the "Poisson bracket", where for some functions f, g their Poisson bracket denoted as $\{f, g\}$ is defined as:

$$\{f, g\} = \sum_i \frac{\partial f}{\partial q}\frac{\partial g}{\partial p_i} - \frac{\partial f}{\partial p_i}\frac{\partial g}{\partial q_i} \tag{269}$$

22. Lecture 22 - Poisson Brackets

For any function $f = f(q_i, p_i, t)$,

$$\frac{\mathrm{d}f}{\mathrm{d}t} = \{f, \mathcal{H}\} + \frac{\partial f}{\partial t} \tag{270}$$

Suppose $f = f(q_i, p_i)$ and $\frac{\partial f}{\partial t} = 0$. Then,

$$\frac{\mathrm{d}f}{\mathrm{d}t} = \{f, \mathcal{H}\} \tag{271}$$

If $\{f, \mathcal{H}\} = 0$, then f is conserved, and f is an integral of motion.

Properties of the Poisson Bracket

1. Antisymmetry $\{f, g\} = -\{g, f\}$
2. Bi-linearity $\{\alpha f + \beta g, h\} = \alpha\{f, h\} + \beta\{g, h\}$ for α, β constants
3. Leibniz rule $\{fg, h\} = f\{g, h\} + \{f, h\}g$
4. Cyclic symmetry (Jacobian Identity) $\{f, \{g, h\}\} + \{g, \{h, f\}\} + \{h, \{f, g\}\} = 0$

Let us look at the Poisson brackets of arbitrary functions. For example,

$$\{q_i, q_j\} = \sum_k \left(\frac{\partial q_i}{\partial q_k} \underbrace{\frac{\partial q_j}{\partial p_k}}_{\text{goes to } 0} - \underbrace{\frac{\partial q_i}{\partial p_k}}_{goes\ to\ 0} \frac{\partial q_i}{\partial q_k} \right) = 0 \tag{272}$$

Those term vanish to zero since the coordinates don't depend on the momenta. For a similar reason, $\{p_i, p_j\} = 0$. Another example:

$$\{q_i, p_j\} = \sum_k \left(\overbrace{\frac{\partial q_i}{\partial q_k}}^{\delta_{ik}} \underbrace{\frac{\partial p_j}{\partial p_k}}_{\delta_{ik}} - \underbrace{\frac{\partial q_i}{\partial p_k} \frac{\partial p_j}{\partial q_k}}_{0} \right) \tag{273}$$

$$= \sum_k \delta_{ik}\delta_{jk} \tag{274}$$

$$= \delta_{ij} \tag{275}$$

This looks like the commutator from quantum mechanics: The position operators commute (they go to 0), the momenta operators commute (they go to 0), but $[\hat{x}_i, \hat{p}_j] = i\hbar\delta_{ij}$. (It turns out in the classical mechanics, the equations of motions in QM gives us Hamilton's equations, though we do have be careful about what we say here).

Let us look at another example: The cyclic coordinates $\left(\frac{\partial \mathcal{H}}{\partial q_i}\right) = 0$.

$$\{p_i, \mathcal{H}\} = \sum_k \left[\frac{\partial p_i}{\partial q_k} \frac{\partial \mathcal{H}}{\partial p_k} - \frac{\partial p_i}{\partial p_k} \underbrace{\frac{\partial \mathcal{H}}{\partial q_k}}_{\text{goes to } 0} \right] \tag{276}$$

$$= 0 \tag{277}$$

So we say p_i and $\mathcal{H}$ "Poisson commute". Suppose we have two integrals of motion I, J. Then,

$$\{\mathcal{H}, I\} = \{\mathcal{H}, J\} = 0 \tag{278}$$

By the Jacobian identity,

$$\{\mathcal{H}, \{I, J\}\} = -\{I, \overbrace{\{J, \mathcal{H}\}}^{0}\} - \{J, \underbrace{\{\mathcal{H}, I\}}_{0}\} \tag{279}$$

$$= 0 \tag{280}$$

This means that the Poisson brackets of I, J is also an integral of motion/conserved. This means the constants of motions form a closed algebra under the Poisson brackets. In mathematics, this thing is called a "Lie algebra".

From now on, the notes will be in index notation (unless the sum is written out explicitly). For example, $\vec{M} = \vec{r} \times \vec{p}$, then $M_i = \sum_{jk} \varepsilon_{ijk} r_j p_k = \varepsilon_{ijk} r_j p_k$ (where repeated indices imply a sum over them, $\varepsilon_{123} = +1, \varepsilon_{213} = -1$, etc. $\varepsilon_{112} = 0$).

Let us suppose the M_x, M_y are conserved. Then immediately, we know that the Poisson bracket of M_x and M_y are conserved. Writing it out:

$$\{M_1, M_2\} = \{r_2 p_3 - r_3 p_2, r_3 p_1 - r_1 p_3\} \tag{281}$$

$$= \{r_2 p_3, r_3 p_1\} + \{r_3 p_2, r_1 p_3\} - \{r_3 p_2, r_3 p_1\} - \{r_2 p_3, r_1 p_3\} \tag{282}$$

The last two terms trivially equal to 0 if you stare at it long enough. The other former two terms can be expanded via Leibniz rule. After many steps of omitted algebra

$$\Rightarrow \{M_1, M_2\} = -p_1 r_2 + r_1 p_2 \tag{283}$$

$$= M_3 \tag{284}$$

We learnt something: We cannot have a system where M_x, M_y is conserved but M_z is not. If M_x, M_y are conserved, then so is M_z. As an exercise, you can show that $\{M^2, M_i\} = 0$, once again similar to something we saw from QM. In fact, the Poisson brackets in classical mechanics are analogous to $i\hbar \times \text{commutator}$ in quantum mechanics, known as the "canonical quantization".

Mathematically, is it possible to design a coordinate transformation of the following form:

$$q_i \to Q_i(q, p), \; p_i \to P_i(q, p) \tag{285}$$

The answer is that we have to be careful, since q, p appear in Hamilton's equations in different signs, but it does exist. These transformations are known as canonical transformations if in these new variables, Hamilton's equations $\dot{Q}_i = \frac{\partial \mathcal{H}}{\partial P_i}$ and $\dot{P}_i = -\frac{\partial \mathcal{H}}{\partial Q_i}$ still hold.

23. Lecture 23 - Canonical Transformations

In the Lagrangian formalism, we naturally chose a coordinate system for a given problem (e.g. Cartesian for simple planar motion, polar for pendulums, etc.) - however, Hamilton's equations are symmetric in p_i's

and q_i's, allowing us to consider a larger class of transformations mentioned at the end of last class, where the new generalized coordinates and functions of q, p.

$$q_i \to Q_i(q,p), \quad p_i \to P_i(q,p) \tag{286}$$

We want to ensure that the transformed coordinate systems obey Hamilton's equations, so the following should still hold:

$$Q_i = \frac{\partial \mathcal{H}}{\partial \dot{P}_i}, \quad P_i = -\frac{\partial \mathcal{H}}{\partial \dot{Q}_i} \tag{287}$$

The class of transformations which satisfy Hamilton's equations are referred to as the **canonical transformations**. Let us determine how it should look like for a system with a single degree of freedom, then generalize this to systems with n degrees of freedom. Given q, p, assume some arbitrary generalized coordinate and generalized momentum $Q(q,p), P(q,p)$.

$$\Rightarrow \dot{Q} = \frac{\partial Q}{\partial q}\dot{q} + \frac{\partial Q}{\partial p}\dot{p} \tag{288}$$

$$\Rightarrow \dot{P} = \frac{\partial P}{\partial q}\dot{q} + \frac{\partial P}{\partial p}\dot{p} \tag{289}$$

Equivalently,

$$\begin{pmatrix} \dot{Q} \\ \dot{P} \end{pmatrix} = \underbrace{\begin{pmatrix} \frac{\partial Q}{\partial q} & \frac{\partial Q}{\partial p} \\ \frac{\partial P}{\partial q} & \frac{\partial P}{\partial p} \end{pmatrix}}_{\mathcal{J}} \begin{pmatrix} \dot{q} \\ \dot{p} \end{pmatrix}$$

The determinant of $\mathcal{J}$ is also known as the Jacobian of the transformation. We can also write down Hamilton's equations in matrix form,

$$\begin{pmatrix} \dot{q} \\ \dot{p} \end{pmatrix} = \underbrace{\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}}_{J} \begin{pmatrix} \frac{\partial \mathcal{H}}{\partial q} \\ \frac{\partial \mathcal{H}}{\partial p} \end{pmatrix} \tag{290}$$

and noting that:

$$\begin{pmatrix} \frac{\partial \mathcal{H}}{\partial q} \\ \frac{\partial \mathcal{H}}{\partial p} \end{pmatrix} = \underbrace{\begin{pmatrix} \frac{\partial Q}{\partial q} & \frac{\partial P}{\partial q} \\ \frac{\partial Q}{\partial p} & \frac{\partial P}{\partial p} \end{pmatrix}}_{\mathcal{J}^T} \begin{pmatrix} \frac{\partial \mathcal{H}}{\partial Q} \\ \frac{\partial \mathcal{H}}{\partial P} \end{pmatrix} \tag{291}$$

Therefore, the new variable obeys:

$$\begin{pmatrix} \dot{Q} \\ \dot{P} \end{pmatrix} = \mathcal{J} J \mathcal{J}^T \begin{pmatrix} \frac{\partial \mathcal{H}}{\partial Q} \\ \frac{\partial \mathcal{H}}{\partial P} \end{pmatrix} \tag{292}$$

This means that if $\mathcal{J} J \mathcal{J}^T = J$ holds, then this is just Hamilton's equations under the new variables:

$$\dot{Q} = \frac{\partial \mathcal{H}}{\partial P}, \quad \dot{P} = -\frac{\partial \mathcal{H}}{\partial Q} \tag{293}$$

and the transformation is canonical. What happens to the Poisson brackets under canonical transformations? Consider $f = f(q,p), g = g(q,p)$:

$$\{f,g\} = \frac{\partial f}{\partial q}\frac{\partial g}{\partial p} - \frac{\partial g}{\partial q}\frac{\partial f}{\partial p} \tag{294}$$

$$= \begin{pmatrix} \frac{\partial f}{\partial q} & \frac{\partial f}{\partial p} \end{pmatrix} \underbrace{\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}}_{J} \begin{pmatrix} \frac{\partial g}{\partial q} \\ \frac{\partial g}{\partial p} \end{pmatrix} \tag{295}$$

and we also know that:

$$\begin{pmatrix} \frac{\partial g}{\partial q} \\ \frac{\partial g}{\partial p} \end{pmatrix} = \underbrace{\begin{pmatrix} \frac{\partial Q}{\partial q} & \frac{\partial P}{\partial q} \\ \frac{\partial Q}{\partial p} & \frac{\partial P}{\partial p} \end{pmatrix}}_{\mathcal{J}^T} \begin{pmatrix} \frac{\partial g}{\partial Q} \\ \frac{\partial Q}{\partial P} \end{pmatrix} \tag{296}$$

and similar for the function f. Putting this into the Poisson bracket, we get:

$$\{f,g\} = \begin{pmatrix} \frac{\partial f}{\partial q} & \frac{\partial f}{\partial p} \end{pmatrix} \underbrace{\mathcal{J}J\mathcal{J}^T}_{=J} \begin{pmatrix} \frac{\partial g}{\partial Q} \\ \frac{\partial Q}{\partial P} \end{pmatrix} \tag{297}$$

$$= \frac{\partial f}{\partial Q}\frac{\partial g}{\partial P} - \frac{\partial g}{\partial Q}\frac{\partial f}{\partial P} \tag{298}$$

$$\tag{299}$$

These are the Poisson brackets calculated with Q and P, so the Poisson brackets are preserved under canonical transformations. In particular, $\{Q,Q\} = \{P,P\} = 0, \{Q,P\} = 1$. We can easily extend this to a system of n variables.

For a system of n degrees of freedom, we have that a transformation $q_i \to Q_i(q,p), p_i \to P_i(q,p)$ is canonical if and only if $\mathcal{J}J\mathcal{J}^T = J$, where:

$$J = \begin{pmatrix} 0_{n\times n} & \mathcal{I}_{n\times n} \\ -\mathcal{I}_{n\times n} & 0_{n\times n} \end{pmatrix} \tag{300}$$

and

$$\mathcal{J} = \begin{pmatrix} \frac{\partial Q_1}{\partial q_1} & \frac{\partial Q_1}{\partial q_2} & \cdots & \frac{\partial Q_1}{\partial q_n} & \frac{\partial Q_1}{\partial p_1} & \frac{\partial Q_1}{\partial p_2} & \cdots & \frac{\partial Q_1}{\partial p_n} \\ \frac{\partial Q_2}{\partial q_1} & \cdots & & & & & & \frac{\partial Q_2}{\partial p_n} \\ \vdots & & & & & & & \vdots \\ \frac{\partial P_n}{\partial q_1} & \cdots & & & & & & \frac{\partial P_n}{\partial p_n} \end{pmatrix} \tag{301}$$

Since the Poisson brackets are conserved under canonical transformations, if we want to verify whether a transformation is canonical, it is sufficient to verify whether:

$$J = \mathcal{J}J\mathcal{J}^T \tag{302}$$

holds, or checking the Poisson brackets:

$$\{Q_i, Q_j\} = \{P_i, P_j\} = 0 \tag{303}$$

$$\{Q_i, P_j\} = \delta_{ij} \tag{304}$$

For example, for a system with one degree of freedom, we can define $Q_i = p_i$ and $P_i = -q_i$, this clearly preserves the Poisson brackets (check for yourself!).

Hence, it is actually a matter of convention of what we call “coordinates" and “momentum" in this picture, since we can always swap them (with a minus sign) using a canonical transformation.

Let us now consider the example of simple harmonic motion:

$$\mathcal{H} = \frac{p^2}{2m} + \underbrace{\frac{1}{2}m\omega^2 q^2}_{=\frac{1}{2}kq^2, \omega = \sqrt{\frac{k}{m}}} \tag{305}$$

Then Hamilton’s equations generate:

$$\dot{p} = -\frac{\partial \mathcal{H}}{\partial q} = -m\omega^2 q \tag{306}$$

$$\dot{q} = \frac{\partial \mathcal{H}}{\partial p} = \frac{p}{m} \tag{307}$$

$$\Rightarrow \ddot{q} = \frac{1}{m}\dot{p} = -\omega^2 q \tag{308}$$

Solving the equations of motion yield:

$$\Rightarrow q(t) = A\sin\left[\omega(t - t_0)\right] \tag{309}$$

$$\Rightarrow p(t) = m\dot{q} = m\omega A\cos\left[\omega(t - t_0)\right] \tag{310}$$

This forms the following closed trajectories (ellipses) in the phase space,

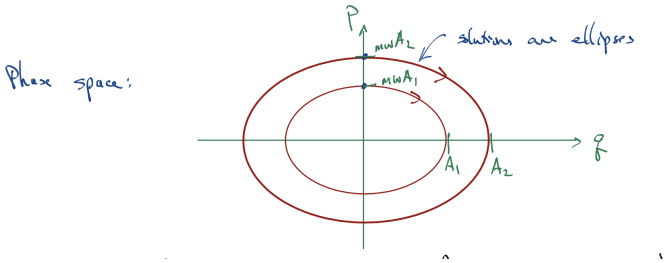


Figure 11: Phase space plot, by Michael Luke [2]

There is a nice canonical transformation to choose for this situation: $Q = \sqrt{m\omega}q, P = p/\sqrt{m\omega}$. This obviously preserves the Poisson brackets, and the Hamiltonian reduces to:

$$\mathcal{H} = \frac{\omega}{2}(P^2 + Q^2) \tag{311}$$

For which the solution becomes:

$$Q(t) = A\sin\left[\omega(t - t_0)\right] \tag{312}$$

$$P(t) = A\cos\left[\omega(t - t_0)\right] \tag{313}$$

We can immediately recognize that these equations forms a circle instead of an ellipse in the phase space of (Q, P). Let us make an astute observation here: For a given orbit (phase space trajectory) in this solution,

the radius of the orbit is constant - this implies that it is an integral of motion!

$$\frac{\mathrm{d}(P^2+Q^2)}{\mathrm{d}t}=0 \tag{314}$$

and since $\mathcal{H} \propto P^2+Q^2$, $\mathcal{H}$ is conserved under these coordinates! Let us a do a canonical transformation to "polar" coordinates (θ, I), not in the real space but in the phase space. θ is our generalized coordinate and I is our generalized momentum. Let:

$$Q=\sqrt{2I}\sin\theta \tag{315}$$

$$P=\sqrt{2I}\cos\theta \tag{316}$$

$$\Rightarrow I=\frac{1}{2}(P^2+Q^2) \tag{317}$$

$$\Rightarrow \theta=\arctan(P/Q) \tag{318}$$

(Verify yourself that this is canonical). Then we have that $\mathcal{H}=I\omega$, which is independent of θ. Therefore, we can immediately recognize that θ is a cyclic variable and its conjugate momentum I is conserved. Explicitly, Hamilton's equations are:

$$\dot{\theta}=\omega, \quad \dot{I}=0 \tag{319}$$

The phase space flow would look like:

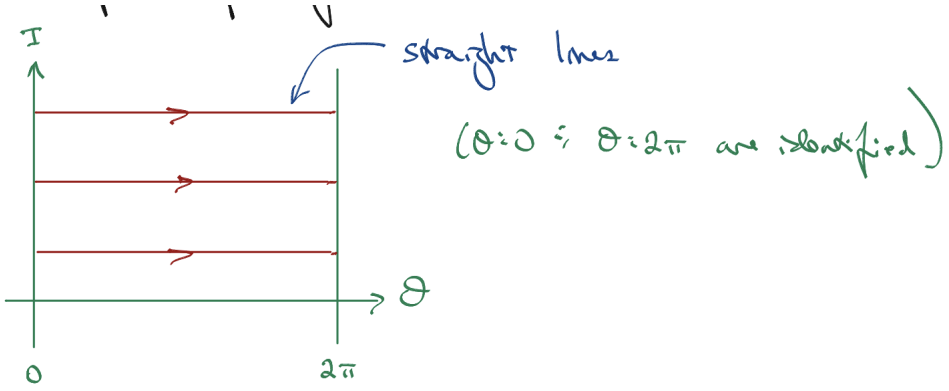


Figure 12: Phase space plot, by Michael Luke [2]

We can see that these are simply constant trajectories independent of θ. (θ, I) are known as the "action-angle" variables, and we can generalize this to more complex systems.

We convert to action-angle variables by finding a canonical transformation such that $\mathcal{H}$ **does not** depend on θ_i's, so they are cyclic. Then we have that:

$$\dot{I}_i=-\frac{\partial \mathcal{H}}{\partial \theta_i}=0 \tag{320}$$

So the I's are conserved, and we have that:

$$\dot{\theta}_i=\frac{\partial \mathcal{H}}{\partial I}=\omega_i(I_1,...,I_n) \tag{321}$$

$$\Rightarrow \theta_i=\omega_i t \tag{322}$$

So the equations of motion are simple in these coordinates. For a general system, we cannot always do this. We will need n conserved quantities so our system is integrable. How do we do this?

Liouville-Arnold Theorem: Any system undergoing bounded motion which has n mutually Poisson commuting integrals of motion $I_1, ..., I_n$ such that $\{I_i, I_J\} = 0$ can be written in action-angle variables and is integrable. (This is often taken as the definition of integrability, also referred to as "Liouville Integrability").

For example, a particle in a central potential has $n = 3$ degrees of freedom. We found 4 (general potential or 5 ($\frac{1}{r}$ potential) integrals of motion $\mathcal{H}, \vec{M}, \vec{A}$ but these do not mutually Poisson commute $\{M_i, M_j\} = \varepsilon_{ijk} M_k$. However, $\mathcal{H}, M^2, M_z$ all Poisson commute, so they are Liouville integrable, and therefore can be written in action-angle variables.

$$I_\varphi = p_\varphi = mr^2\dot{\varphi} \tag{323}$$

$$I_r = \sqrt{\frac{m}{2|E|}}k - I_\varphi \tag{324}$$

$$\Rightarrow \mathcal{H}(I_r, I_\varphi) = -\frac{mk^2}{(I_r + I_\varphi)^2} \tag{325}$$

Note that $\omega_r = \frac{\partial \mathcal{H}}{\partial I_r} = \frac{\partial \mathcal{H}}{\partial I_\varphi} = \omega_\varphi$ where both θ_r, θ_φ frequencies are the same, giving closed orbits.

24. Lecture 24 - Conservation Laws II

References

[1] L. D. Landau and E. M. Lifshitz. *Mechanics, Third Edition: Volume 1 (Course of Theoretical Physics)*. 3rd ed. Butterworth-Heinemann, Jan. 1976. ISBN: 0750628960. URL: http://www.worldcat.org/isbn/0750628960.

[2] Michael Luke. *PHY354-2024 III 25-54*. University of Toronto, Department of Physics. 2024.

[3] David Morin. *Introduction to Classical Mechanics: With Problems and Solutions*. Cambridge University Press, 2008.

[4] David Tong. *Classical Dynamics*. University of Cambridge, Department of Applied Mathematics and Theoretical Physics. 2015.