

Fundamentals of Classical Mechanics

ADVANCED FORMALISMS FROM VARIATIONAL PRINCIPLES

Hasanat Hasan

CUNY Graduate Center
Physics PhD Track
`hhasan@gradcenter.cuny.edu`

January 14, 2026

Preface

The study of Classical Mechanics is often glossed over and unappreciated with the students' priority on quantum mechanics. However, students can still derive great appreciation for mechanics by understanding how the founders developed quantum theory with only the tools and mastery in classical mechanics. We aim to highlight the immense importance of variational principles in physics, in fact, all sections in this document are a consequence of one. We cover foundational topics such as Lagrangian and Hamiltonian formulations, Noether's Theorem, Canonical Transformations, Hamilton-Jacobi Theory, and Action-Angle Variables.

Hasanat Hasan
New York, NY

Contents

1	Lagrangian Mechanics	3
1.1	Derivation	3
1.2	Euler-Lagrange Equations	4
2	Hamiltonian Mechanics	5
2.1	Hamilton's Derivation	5
2.2	Hamilton's Equations from The Euler-Lagrange Eqns.	6
2.2.1	Time Evolution	7
3	Noether's Theorem	8
3.1	Derivation	8
3.2	Examples	9
3.2.1	Translational Symmetry and Linear Momentum	9
3.2.2	Rotational Symmetry and Angular Momentum	9
3.2.3	Time Translational Symmetry	10
3.2.4	Time Translational Symmetry and Energy Conservation	11
4	Canonical Transformations	12
4.1	Derivation	12
4.2	Infinitesimal Canonical Transformations	13
4.3	The Hamiltonian and Infinitesimal Time Evolution	13
4.4	The Action and Time Evolution	14
5	Hamilton-Jacobi Theory	16
5.1	Derivation	16
5.2	Time Independent HJE (TIHJE)	17
5.2.1	Example with Harmonic Oscillator	18
5.3	Application to The Kepler Problem	19
6	Action-Angle Variables	21
6.1	Derivation	21
6.1.1	Frequency of Oscillation	22
6.1.2	Harmonic Oscillator Revisited	23
6.1.3	Quartic Oscillator Example	23
6.1.4	Kepler Orbits Revisited	24
6.2	Canonical Perturbation Theory	25
6.2.1	Perturbed Oscillator	27
6.2.2	Simple Pendulum	28
6.2.3	Perturbed Kepler Orbits	29
6.3	Adiabatic Invariance	30
6.3.1	Adiabatic Pendulum	30
6.3.2	Adiabatic Well	31
A	Calculus Review	33
A.1	The Chain Rule for Change of Variables	33
A.2	Differentials of Multivariable Functions	33
A.3	Differentiation Under the Integral Sign (Leibniz Rule)	33
B	Calculus of Variations	33
C	Alternative Derivations:	34
C.1	Hamilton's Equations	34
D	Useful Resources	35

1 Lagrangian Mechanics

1.1 Derivation

We shall introduce Lagrangian mechanics—an extremely powerful reformulation of mechanics not in terms of force vectors, but in terms of energies and generalized coordinates. We begin by applying [d’Lambert’s Principle of virtual work](#) to Newton’s second law:

$$\sum_i (\vec{F}_i - m_i \ddot{\vec{q}}_i) \cdot \delta \vec{q}_i = 0 \quad (1)$$

Where we project the difference of the forces and $m\ddot{q}$ along infinitesimal displacements $\delta \vec{q}$. This is called the virtual work. Notice that Eq.(1) states that the work along neighboring paths $\delta \vec{q}_i$ is zero meaning it is stable in neighboring paths. also note that the variation at the end points must be zero, because we only know the end points but not the path by the object. If we assume all forces are conservative then there exists a scalar field for which the force is the gradient of (gradient theorem). Since we are considering infinitesimal variations, the force-potential relation becomes:

$$\vec{F} = -\nabla U \rightarrow -\delta U$$

So Eq.(1) becomes:

$$\begin{aligned} \sum_i (\vec{F}_i - m_i \ddot{\vec{q}}_i) \cdot \delta \vec{q}_i &= 0 \\ \sum_i \vec{F}_i \cdot \delta \vec{q}_i &= \sum_i m_i \ddot{\vec{q}}_i \cdot \delta \vec{q}_i \\ -\sum_i \delta U_i &= \sum_i m_i \ddot{\vec{q}}_i \cdot \delta \vec{q}_i \end{aligned}$$

We can apply the following identity: $\ddot{q}\delta q = \frac{d}{dt}(\dot{q}\delta q) - \dot{q}\delta\dot{q} = \frac{d}{dt}(\dot{q}\delta q) - \frac{1}{2}\delta(\dot{q}^2)$ (by chain rule).

$$\begin{aligned} -\sum_i \delta U_i &= \sum_i m_i \ddot{\vec{q}}_i \cdot \delta \vec{q}_i \\ &= \sum_i m_i \frac{d}{dt}(\dot{\vec{q}}_i \cdot \delta \vec{q}_i) - \frac{1}{2} \sum_i m_i \delta(|\dot{\vec{q}}_i|^2) \end{aligned}$$

Notice that $\sum_i \delta U_i$ is just the total variation of the potential energy, lets denote it as δU . and lets do the same for the $\frac{1}{2} \sum_i m_i \delta(|\dot{\vec{q}}_i|^2)$ which is the sum of the variation of the kinetic energies, which is δK . substituting these back in:

$$\begin{aligned} -\delta U &= \sum_i m_i \frac{d}{dt}(\dot{\vec{q}}_i \cdot \delta \vec{q}_i) - \delta K \\ \delta(K - U) &= \sum_i m_i \frac{d}{dt}(\dot{\vec{q}}_i \cdot \delta \vec{q}_i) \\ \int_a^b dt \delta(K - U) &= \int_a^b dt \sum_i m_i \frac{d}{dt}(\dot{\vec{q}}_i \cdot \delta \vec{q}_i) \\ \delta \int_a^b dt (K - U) &= \sum_i m_i \frac{d}{dt}(\dot{\vec{q}}_i \cdot \delta \vec{q}_i) \Big|_a^b \end{aligned}$$

Recall that the variation at the endpoints must be 0:

$$\boxed{\delta \int_a^b (K - U) dt = 0} \quad (2)$$

THIS is Hamilton’s principle, by using d’Lambert’s principle, we showed that the variation of the displacements along neighboring paths must be zero (or stationary) if it satisfies Eq. (2). Now we

have reason to place importance on the quantity $K - U$ as it is the function that must be stationary to satisfy Hamilton's principle, we define it as Lagrangian:

$$L = K - U \quad (3)$$

and we call its integral the action:

$$S = \int_a^b L[q, \dot{q}, t] dt \quad (4)$$

Through this derivation, we have shown that

- The variation of the action must be 0 (Eq. (2))
- The origin of the Lagrangian definition (Eq. (3))
- The origin of the action definition (Eq. (4))

all as a result of d'Lambert's principle. This derivation was found from chapter 2 of [Arnold Sommerfeld's Mechanics](#), [Dr. Jorge Diaz's](#) video on this derivation. and [chapter 1.4 of Goldstein, Safko, and Poole's Mechanics](#)

1.2 Euler-Lagrange Equations

Given the Lagrangian, $L(q(t), \dot{q}(t), t)$, with fixed end points ($\delta q(a) = \delta q(b) = 0$), we apply Hamilton's Principle by extremizing the action to find the dynamics of a physical system given by $q(t)$. The action is given by

$$S[q] = \int_a^b L(q(t), \dot{q}(t), t) dt$$

We make use of the variational operator δ , and extremize $S[q]$ by $\delta S = 0$:

$$\begin{aligned} \delta S &= \delta \int_a^b L(q(t), \dot{q}(t), t) dt = 0 \\ &= \int_a^b \delta L(q(t), \dot{q}(t), t) dt = 0 \end{aligned}$$

Here we make use of the fact that the variation of a function, $f(x_i(t))$ is given by:

$$\delta f(x_i(t)) = \sum_i \delta x_i(t) \frac{\partial f}{\partial x_i}$$

Applying this to the Lagrangian:

$$\delta S = \int_a^b \left(\delta q \frac{\partial L}{\partial q} + \delta \dot{q} \frac{\partial L}{\partial \dot{q}} \right) dt = 0$$

Next we integrate the second term by parts to obtain a δq term: Let $u = \frac{\partial L}{\partial \dot{q}} \rightarrow du = \frac{d}{dt} \frac{\partial L}{\partial \dot{q}} dt$, and $dv = \frac{d}{dt} (\delta q) dt \rightarrow v = \delta q$

$$\begin{aligned} \int_a^b \delta \dot{q} \frac{\partial L}{\partial \dot{q}} dt &= \left(\frac{\partial L}{\partial \dot{q}} \right) \delta q \Big|_a^b - \int_a^b \delta q \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}} \right) dt \\ &= 0 - \int_a^b \delta q \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}} \right) dt \end{aligned}$$

Substituting back into the action:

$$\delta S = \int_a^b dt \left(\frac{\partial L}{\partial q} - \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}} \right) \right) \delta q = 0$$

For any variation, $\delta q(t)$, $\delta S = 0$ only if:

$$\boxed{\frac{\partial L}{\partial q} - \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}} \right) = 0} \quad (5)$$

This is the Euler-Lagrange equation.

2 Hamiltonian Mechanics

2.1 Hamilton's Derivation

In Lagrangian mechanics, the physical trajectory is the one that makes the action stationary (Hamilton's principle). The physical path is given by the solution to the Euler-Lagrange equation in Eq.(5), which is a 2nd order partial differential equation. Hamilton wondered if there was a way to convert this into 2 first order differential equations, as they would be much easier to solve. His exploration led to another reformulation of mechanics in terms of the "Hamiltonian". We first consider the kinetic energy K as function of (q, \dot{q}) in the form

$$K(q, \dot{q}) = \frac{1}{2} \sum_i m_i \dot{q}_i^2$$

with differential $dK = \sum_i \left(dq_i \frac{\partial K}{\partial q_i} + d\dot{q}_i \frac{\partial K}{\partial \dot{q}_i} \right)$. Differentiating K wrt \dot{q}_j we obtain $\frac{\partial K}{\partial \dot{q}_j} = m_j \dot{q}_j$. We can substitute this back into the kinetic energy to obtain $2K = \sum_i \frac{\partial K}{\partial \dot{q}_i} \dot{q}_i$. From the expression derived, the differential of $2K$ is given by $2dK = \sum_i \left(d\dot{q}_i \frac{\partial K}{\partial \dot{q}_i} + d \left(\frac{\partial K}{\partial \dot{q}_i} \right) \dot{q}_i \right)$. Taking the difference of $2dK - dK$ we get:

$$\begin{aligned} 2dK - dK &= \sum_i \left(d\dot{q}_i \frac{\partial K}{\partial \dot{q}_i} + d \left(\frac{\partial K}{\partial \dot{q}_i} \right) \dot{q}_i - dq_i \frac{\partial K}{\partial q_i} - d\dot{q}_i \frac{\partial K}{\partial \dot{q}_i} \right) \\ dK &= \sum_i \left(d \left(\frac{\partial K}{\partial \dot{q}_i} \right) \dot{q}_i - dq_i \frac{\partial K}{\partial q_i} \right) \end{aligned}$$

Now we introduce the canonical momentum $p_i = \frac{\partial K}{\partial \dot{q}_i}$, which is independent of q . p is in general a function of \dot{q} so $p \equiv p(\dot{q})$. Hamilton sought to invert this relation to obtain $\dot{q} = \dot{q}(p)$ and substitute into $K(q, \dot{q}) : K(q, \dot{q}) \rightarrow K(q, \dot{q}(p)) \rightarrow \tilde{K}(q, p) = K(q, \dot{q}(p))$. Now we apply the definition $\dot{q} = \dot{q}(p)$ and $p_i = \frac{\partial K}{\partial \dot{q}_i}$ to dK to obtain

$$d\tilde{K} = \sum_i \left(dp_i \dot{q}_i - dq_i \frac{\partial K}{\partial q_i} \right)$$

Separately, the definition of $d\tilde{K}(q, p)$ is given by $d\tilde{K}(q, p) = \sum_i \left(dq_i \frac{\partial \tilde{K}}{\partial q_i} + dp_i \frac{\partial \tilde{K}}{\partial p_i} \right)$, equating the differentials we obtain:

$$\dot{q}_i = \frac{\partial \tilde{K}}{\partial p_i}, \quad \frac{\partial K}{\partial q_i} = -\frac{\partial \tilde{K}}{\partial q_i}$$

Focusing on the $\dot{q}_i = \frac{\partial \tilde{K}}{\partial p_i}$ equation, we assume the potential is independent of \dot{q} and given by $V = V(q, t)$, so $\frac{\partial V}{\partial p} = 0$. This allows us to add $\frac{\partial V}{\partial p}$ to \dot{q} :

$$\begin{aligned} \dot{q}_i + 0 &= \frac{\partial \tilde{K}}{\partial p_i} + \frac{\partial V}{\partial p} = \frac{\partial}{\partial p} (\tilde{K} + V) \dot{q}_i + 0 \\ &= \frac{\partial \tilde{K}}{\partial p_i} + \frac{\partial V}{\partial p} \\ \dot{q}_i &= \frac{\partial}{\partial p} (\tilde{K} + V) \end{aligned}$$

Recall that the Lagrangian is given by $L(q, \dot{q}, t) = K(q, \dot{q}) - V(q, t)$ (assuming the kinetic energy is of the same form and the potential is independent of \dot{q}). Then $\frac{\partial L}{\partial \dot{q}_i} = \frac{\partial K}{\partial \dot{q}_i} - 0$, which means $p_i = \frac{\partial K}{\partial \dot{q}_i} = \frac{\partial L}{\partial \dot{q}_i}$. We can rewrite the Euler-Lagrange equations from Eq.(5) using p_i to get $\dot{p}_i = \frac{\partial}{\partial q_i} (K - V)$. Then using

the derived relation $\frac{\partial K}{\partial q_i} = -\frac{\partial \tilde{K}}{\partial q_i}$ we obtain:

$$\begin{aligned}\dot{p}_i &= \frac{\partial}{\partial q_i} (K - V) = \frac{\partial K}{\partial q_i} - \frac{\partial V}{\partial q_i} \\ &= -\frac{\partial \tilde{K}}{\partial q_i} - \frac{\partial V}{\partial q_i} \\ \dot{p}_i &= -\frac{\partial}{\partial p} (\tilde{K} + V)\end{aligned}$$

We once again see $\tilde{K} + V$, we now have a reason to promote this to a definition, let $H \equiv \tilde{K} + V$, then the equations for \dot{q}_i and \dot{p}_i become

$$\dot{q}_i = \frac{\partial H}{\partial p}, \quad \dot{p}_i = -\frac{\partial H}{\partial p}$$

These are Hamilton's canonical equations. We still need some form of a "Hamiltonian action" to rederive from Hamilton's principle, which requires a direct relation between H and L . We hence replace the \dot{q} in L to obtain $L(q, \dot{q}, t) = K(q, \dot{q}) - V(q, t) \rightarrow \tilde{L}(q, \dot{q}(p), t) = \tilde{K}(q, p) - V(q, t)$. If we add $\tilde{L} + H$, we get:

$$\begin{aligned}\tilde{L} + H &= \tilde{K} - V + \tilde{K} + V \\ &= 2\tilde{K}\end{aligned}$$

Recall that $2K = \sum_i \frac{\partial K}{\partial \dot{q}_i} \dot{q}_i$, which becomes $2\tilde{K} = \sum_i p_i \dot{q}(p_i)$, the equation becomes

$$\begin{aligned}\tilde{L} + H &= 2\tilde{K} = \sum_i p_i \dot{q}(p_i) \\ H(q, p, t) &= \sum_i p_i \dot{q}(p_i) - \tilde{L}(q, \dot{q}(p), t)\end{aligned}$$

$$\boxed{H(q_i, p_i, t) = \sum p \dot{q} - L} \quad (6)$$

Now that we have a direct relation to L , the action can extended in terms of (q, p) : $S[q] = \int L dt \rightarrow S[q, p] = \int (\sum p \dot{q} - L) dt$, in the next section, we show we can derive Hamilton's equations from extremizing the action to show its consistency with Hamilton's principle.

2.2 Hamilton's Equations from The Euler-Lagrange Eqns.

Previously we applied Hamilton's Principle to the Lagrangian to find a path through configuration space that minimizes the action, and found it must be a solution to Eq. 5. Following in the same spirit, we shall derive Hamilton's equations through the modified Hamilton's principle: instead of variations of a path through configuration space (q, \dot{q}) , we consider variations of paths through phase space (q, p) . The modified Hamilton's principle states that the true path an object follows is one that minimizes the action over variations through phase space:

$$\delta S = \delta \int_a^b dt F(q, \dot{q}; p, \dot{p}, t) = 0$$

The Lagrangian is given by $L = \sum_i p_i \dot{q}_i - H(q, p, t)$, and by the modified Hamilton's principle:

$$\delta S = \delta \int_a^b (\sum_i p_i \dot{q}_i - H(q_i, p_i, t)) dt$$

The integral is of the form $F(q, \dot{q}; p, \dot{p}, t) = \sum_i p_i \dot{q}_i - H(q_i, p_i, t)$. Applying the Euler-Lagrange equations:

$$\begin{aligned}\frac{d}{dt} \left(\frac{\partial F}{\partial \dot{q}_i} \right) &= \frac{\partial F}{\partial q_i} \\ \frac{d}{dt} \left(\frac{\partial F}{\partial \dot{p}_i} \right) &= \frac{\partial F}{\partial p_i}\end{aligned}$$

Evaluating for the q_i variables:

$$\begin{aligned} \rightarrow \frac{\partial F}{\partial \dot{q}_i} &= \frac{\partial}{\partial \dot{q}_i} \left(\sum_j p_j \dot{q}_j - H(q_j, p_j, t) \right) = p_i \\ \rightarrow \frac{\partial F}{\partial q_i} &= \frac{\partial}{\partial q_i} \left(\sum_j p_j \dot{q}_j - H(q_j, p_j, t) \right) = -\frac{\partial H}{\partial q_i} \end{aligned}$$

So the Euler-Lagrange eqs become:

$$\frac{d}{dt}(p_i) = -\frac{\partial H}{\partial q_i} \rightarrow \dot{p}_i = -\frac{\partial H}{\partial q_i}$$

Now evaluating for the p_i variables:

$$\begin{aligned} \rightarrow \frac{\partial F}{\partial \dot{p}_i} &= \frac{\partial}{\partial \dot{p}_i} \left(\sum_j p_j \dot{q}_j - H(q_j, p_j, t) \right) = 0 \\ \rightarrow \frac{\partial F}{\partial p_i} &= \frac{\partial}{\partial p_i} \left(\sum_j p_j \dot{q}_j - H(q_j, p_j, t) \right) = \dot{q}_i - \frac{\partial H}{\partial p_i} \end{aligned}$$

Substituting into the Euler-Lagrange eqs:

$$\frac{d}{dt}(0) = \dot{q}_i - \frac{\partial H}{\partial p_i} \rightarrow \dot{q}_i = \frac{\partial H}{\partial p_i}$$

Thus we have derived Hamilton's equations by using the Euler-Lagrange equations through the modified Hamilton's principle for phase space:

$$\boxed{\dot{q}_i = \frac{\partial H}{\partial p_i}, -\dot{p}_i = \frac{\partial H}{\partial q_i}} \quad (7)$$

This derivation is found in Goldstein's mechanics in section 5, chapter 8. An alternative derivation is found in appendix C.1

2.2.1 Time Evolution

Consider the time evolution of the Lagrangian:

$$\frac{d}{dt}L(q_i, \dot{q}_i, t) = \frac{\partial L}{\partial t} + \sum_i \left(\frac{\partial L}{\partial q_i} \frac{dq_i}{dt} + \frac{\partial L}{\partial \dot{q}_i} \frac{d\dot{q}_i}{dt} \right)$$

Recall the Euler-Lagrange equations: $\frac{\partial L}{\partial q} = \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}} \right)$, and $\dot{q}_i = \frac{dq_i}{dt}$, $\ddot{q}_i = \frac{d\dot{q}_i}{dt}$ so we can substitute this into the previous results:

$$\frac{d}{dt}L(q_i, \dot{q}_i, t) = \frac{\partial L}{\partial t} + \sum_i \left(\dot{q}_i \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_i} \right) + \ddot{q}_i \frac{\partial L}{\partial \ddot{q}_i} \right)$$

Next we apply the following mathematical trick using the product rule:

$$\frac{d}{dt} \left(\dot{q}_i \frac{\partial L}{\partial \dot{q}_i} \right) = \ddot{q}_i \frac{\partial L}{\partial \ddot{q}_i} + \dot{q}_i \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_i} \right)$$

Substituting:

$$\begin{aligned} \frac{d}{dt}L(q_i, \dot{q}_i, t) &= \frac{\partial L}{\partial t} + \sum_i \frac{d}{dt} \left(\dot{q}_i \frac{\partial L}{\partial \dot{q}_i} \right) \\ &= \frac{\partial L}{\partial t} + \sum_i \frac{d}{dt} (p_i \dot{q}_i) \end{aligned}$$

$$\begin{aligned}
\sum_i \frac{d}{dt} (p_i \dot{q}_i) - \frac{d}{dt} L(q_i, \dot{q}_i, t) &= -\frac{\partial L}{\partial t} \\
\frac{d}{dt} \left(\sum_i p_i \dot{q}_i - L(q_i, \dot{q}_i, t) \right) &= -\frac{\partial L}{\partial t} \\
\frac{dH}{dt} &= -\frac{\partial L}{\partial t}
\end{aligned}$$

Notice that if the Lagrangian has no explicit time dependence then the Hamiltonian is conserved (constant with time). We show this means energy is conserved and that the Hamiltonian represents the total energy if it is conserved.

3 Noether's Theorem

3.1 Derivation

Theorem: If a continuous, infinitesimal transformation of the coordinates leaves the variation of the Action invariant, that transformation is a Noether symmetry if the Lagrangian changes by at most a total time derivative. For every such symmetry, there is a corresponding conserved quantity

Proof: We begin with a general continuous, infinitesimal, transformation of the coordinates along a physical path (meaning it satisfies the Euler-Lagrange equations.), which is given by $q \rightarrow q' \equiv q + \epsilon \eta(t)$, so $\delta q(t) = \epsilon \eta(t)$, $\delta \dot{q}(t) = \epsilon \dot{\eta}(t)$. The change in the Lagrangian due to this transformation is given by:

$$\begin{aligned}
\delta L &= \delta q \frac{\partial L}{\partial q} + \delta \dot{q} \frac{\partial L}{\partial \dot{q}} + 0 = (\epsilon \eta) \frac{\partial L}{\partial q} + (\epsilon \dot{\eta}) \frac{\partial L}{\partial \dot{q}} \\
&= (\epsilon \eta) \frac{d}{dt} \frac{\partial L}{\partial \dot{q}} + \frac{\partial L}{\partial \dot{q}} \frac{d(\epsilon \eta)}{dt} = \frac{d}{dt} \left(\epsilon \eta \frac{\partial L}{\partial \dot{q}} \right) \\
&= \frac{d}{dt} (p \delta q)
\end{aligned}$$

Where we have applied the Euler-Lagrange equation because this is a variation along a physical path. We made use of the canonical momentum $p = \frac{\partial L}{\partial \dot{q}}$ and $\delta q(t) = \epsilon \eta(t)$. We make note that a general continuous transformation of the coordinates induces a change in the Lagrangian given by

$$\delta L = \frac{d}{dt} (p \delta q)$$

Independently, by Hamilton's Principle, we are allowed to transform the Lagrangian by a first derivative, $\frac{d}{dt}(g(q, t))$, because it leaves the variation of the action invariant, so it yields the same equations of motion:

$$\begin{aligned}
\delta S &= 0 \\
\delta S' &= \int_a^b dt \delta L' = \int_a^b dt \delta \left(L + \frac{d}{dt} (g(q, t)) \right) \\
&= \delta S + \delta g(q, t)|_a^b = \delta S + \frac{\partial g}{\partial q} \delta q(t)|_a^b \\
&= \delta S + \frac{\partial g}{\partial q} [\delta q(b) - \delta q(a)] = \delta S + \frac{\partial g}{\partial q} [0 - 0] \\
&= \delta S = 0
\end{aligned} \tag{8}$$

Therefore the transformed Lagrangian is given by $L' = L + \frac{dg}{dt}$, and the change in the Lagrangian is $\delta L = \frac{dg}{dt}$. This is the general condition for a symmetry of the action.

Equating the two expressions for the change in the Lagrangian, we obtain:

$$\begin{aligned}\frac{d}{dt}(p\delta q) &= \frac{dg}{dt} \\ \frac{d}{dt}(p\delta q - g) &= 0 \\ p\delta q - g &= \text{const.}\end{aligned}\tag{9}$$

Therefore we proven the following profound result, given a general continuous transformation of the coordinates, the Lagrangian changes by $\delta L = \frac{d}{dt}(p\delta q)$, if it changes the Lagrangian by at most a total time derivative $\frac{dg}{dt}$, it must be a symmetry of the action with a corresponding conserved quantity $p\delta q - g$.

3.2 Examples

3.2.1 Translational Symmetry and Linear Momentum

A free particle under no potential has a Lagrangian of the form $L = \frac{1}{2}m\dot{q}^2$. If we apply a translation to the coordinates: $q \rightarrow q' \equiv q + \epsilon$, $\delta q = \epsilon$, $\delta \dot{q} = 0$. The change in the Lagrangian is:

$$\begin{aligned}\delta L &= \delta q \frac{\partial L}{\partial q} + \delta \dot{q} \frac{\partial L}{\partial \dot{q}} \rightarrow \epsilon(0) + 0(m\dot{q}) = 0 \\ \delta L &= 0\end{aligned}$$

This implies that $\delta L = \frac{dg}{dt} = 0 \rightarrow g(q, t) = \text{const.}$ Using Eq.(9), the conserved quantity is: $p\delta q - g = \text{const.} \rightarrow p\epsilon = \text{const.} \rightarrow p = \text{const.}$ Thus arriving at the familiar result that if the Lagrangian is independent of position then the momentum is conserved. In other words, if a translation is a symmetry of the action, then linear momentum is conserved.

Another angle at this is that if we didn't know the form of the Lagrangian, then: $\delta L = \delta q \frac{\partial L}{\partial q} + \delta \dot{q} \frac{\partial L}{\partial \dot{q}} \rightarrow \delta L = \epsilon \frac{\partial L}{\partial q}$. So:

$$\begin{aligned}\delta L &= \epsilon \frac{\partial L}{\partial q} = \frac{dg}{dt} \rightarrow \epsilon \frac{dp}{dt} - \frac{dg}{dt} = 0 \\ &\rightarrow \epsilon p(t) - g(q, t) = \text{const.}\end{aligned}$$

For momentum to be conserved, $g(q, t)$ must be constant, which necessitates $\delta L = \frac{dg}{dt} = 0 \rightarrow g(q, t) = \text{const.}$ Therefore $\delta L = \epsilon \frac{\partial L}{\partial q} = 0$; the Lagrangian must be position independent.

3.2.2 Rotational Symmetry and Angular Momentum

A free particle under no potential has a Lagrangian of the form $L = \frac{1}{2}m\dot{\vec{r}}^2$. If we apply a rotation to the coordinates: $\vec{r} \rightarrow \vec{r}' \equiv \vec{r} + \epsilon(\hat{n} \times \vec{r})$, $\delta \vec{r} = \epsilon(\hat{n} \times \vec{r})$, $\delta \dot{\vec{r}} = \epsilon(\hat{n} \times \dot{\vec{r}})$. The change in the Lagrangian is:

$$\begin{aligned}\delta L &= \delta \vec{r} \cdot \frac{\partial L}{\partial \vec{r}} + \delta \dot{\vec{r}} \cdot \frac{\partial L}{\partial \dot{\vec{r}}} \rightarrow \epsilon(\hat{n} \times \vec{r}) \cdot (0) + \epsilon(\hat{n} \times \dot{\vec{r}}) \cdot (m\dot{\vec{r}}) \\ &= 0 + m\epsilon \hat{n} \cdot (\dot{\vec{r}} \times \dot{\vec{r}}) = 0 \\ \delta L &= 0\end{aligned}$$

This implies that $\delta L = \frac{dg}{dt} = 0 \rightarrow g(\vec{r}, t) = \text{const.}$ Using Eq.(9), the conserved quantity is: $\vec{p} \cdot \delta \vec{r} - g = \text{const.} \rightarrow \vec{p} \cdot [\epsilon(\hat{n} \times \vec{r})] = \text{const.}$ Using the vector identity $\vec{A} \cdot (\vec{B} \times \vec{C}) = (\vec{C} \times \vec{A}) \cdot \vec{B}$, we rearrange to find: $\epsilon \hat{n} \cdot (\vec{r} \times \vec{p}) = \text{const.}$ Thus arriving at the familiar result that if the Lagrangian is rotationally symmetric (scalar), then the angular momentum $\vec{L} = \vec{r} \times \vec{p}$ (specifically the component along the axis of rotation \hat{n}) is conserved.

If we didn't know the form of the Lagrangian, then:

$$\begin{aligned}
\delta L &= \delta \vec{r} \cdot \dot{\vec{p}} + \delta \dot{\vec{r}} \cdot \vec{p} \rightarrow \epsilon(\hat{n} \times \vec{r}) \cdot \dot{\vec{p}} + \epsilon(\hat{n} \times \dot{\vec{r}}) \cdot \vec{p} \\
&= \epsilon \hat{n} \cdot (\vec{r} \times \dot{\vec{p}}) + \epsilon \hat{n} \cdot (\dot{\vec{r}} \times \vec{p}) \\
&= \epsilon \hat{n} \cdot \frac{d}{dt}(\vec{r} \times \vec{p})
\end{aligned}$$

So:

$$\begin{aligned}
\delta L &= \epsilon \hat{n} \cdot \frac{d(\vec{r} \times \vec{p})}{dt} = \frac{dg}{dt} \\
&\rightarrow \epsilon \hat{n} \cdot \frac{d(\vec{r} \times \vec{p})}{dt} - \frac{dg}{dt} = 0 \\
&\rightarrow \epsilon \hat{n} \cdot (\vec{r} \times \vec{p}) - g(\vec{r}, t) = \text{const.}
\end{aligned}$$

For angular momentum to be conserved, $g(\vec{r}, t)$ must be constant, which necessitates $\delta L = \frac{dg}{dt} = 0$. Therefore, the Lagrangian must be invariant under rotation (a scalar quantity depending only on magnitudes $|\vec{r}|$ and $|\dot{\vec{r}}|$).

3.2.3 Time Translational Symmetry

Now we treat time as a dynamical variable to apply time translations. If we parametrize time, t , with τ , such that the new "Extended Lagrangian" is $\mathcal{L} = L(q, \dot{q}, t) \frac{dt}{d\tau}$. We define the following; for time $t = T(\tau)$, $\delta t = \delta T$. For position $q = Q(\tau)$, $\delta q = \delta Q$. The action becomes:

$$\begin{aligned}
S &= \int_a^b L[q, \dot{q}, t] dt = \int_a^b L[q, \dot{q}, t] \frac{dt}{d\tau} d\tau \\
&= \int_a^b \mathcal{L}(Q, Q', T, T') d\tau
\end{aligned}$$

Let $T' = \frac{dt}{d\tau}$, $\mathcal{L} = LT'$. Noting $d(t(\tau)) = d\tau \frac{dt}{d\tau}$, we can substitute this into $\dot{q} = \frac{dq}{dt} = \frac{dq}{d\tau} \frac{d\tau}{dt} = \frac{Q'}{T'}$. If an infinitesimal time translation leaves the action invariant, then:

$$\delta \mathcal{L} = \delta Q \frac{\partial \mathcal{L}}{\partial Q} + \delta Q' \frac{\partial \mathcal{L}}{\partial Q'} + \delta T \frac{\partial \mathcal{L}}{\partial T} + \delta T' \frac{\partial \mathcal{L}}{\partial T'}$$

Using the Euler-Lagrange equations for the independent variables Q, T:

$$\begin{aligned}
\delta \mathcal{L} &= \delta Q \frac{d}{d\tau} \frac{\partial \mathcal{L}}{\partial Q'} + \delta Q' \frac{\partial \mathcal{L}}{\partial Q'} + \delta T \frac{d}{d\tau} \frac{\partial \mathcal{L}}{\partial T'} + \delta T' \frac{\partial \mathcal{L}}{\partial T'} \\
&= \frac{d}{d\tau} \left(\delta Q \frac{\partial \mathcal{L}}{\partial Q'} + \delta T \frac{\partial \mathcal{L}}{\partial T'} \right)
\end{aligned}$$

Evaluating the following partial derivatives using the chain rule: $\frac{\partial \mathcal{L}}{\partial \alpha} = \sum_i \frac{\partial \mathcal{L}}{\partial q_i} \frac{\partial q_i}{\partial \alpha}$ For Q' :

$$\begin{aligned}
\frac{\partial \mathcal{L}}{\partial Q'} &= \frac{\partial(LT')}{\partial Q'} \rightarrow T' \frac{\partial L}{\partial Q'} \rightarrow T' \frac{\partial L}{\partial \dot{q}} \frac{\partial \dot{q}}{\partial Q'} \rightarrow T' \frac{\partial L}{\partial \dot{q}} \left(\frac{1}{T'} \right) \\
\frac{\partial \mathcal{L}}{\partial Q'} &= \frac{\partial L}{\partial \dot{q}}
\end{aligned}$$

For T' :

$$\begin{aligned}
\frac{\partial \mathcal{L}}{\partial T'} &= T' \frac{\partial L}{\partial T'} + L \rightarrow L + T' \left(-\frac{\dot{q}}{T'} \frac{\partial L}{\partial \dot{q}} \right) \\
\frac{\partial \mathcal{L}}{\partial T'} &= L - \dot{q} \frac{\partial L}{\partial \dot{q}}
\end{aligned}$$

Recall $\delta q = \delta Q$, $\delta t = \delta T$, $dt = d\tau \frac{dt}{d\tau} = T' d\tau$, substituting all the results into $\delta \mathcal{L}$:

$$\begin{aligned}\delta \mathcal{L} &= \frac{d}{d\tau} \left(\delta q \frac{\partial L}{\partial \dot{q}} + \delta t \left(L - \dot{q} \frac{\partial L}{\partial \dot{q}} \right) \right) \\ &= \frac{d}{dt} \left(\delta q \frac{\partial L}{\partial \dot{q}} + \delta t \left(L - \dot{q} \frac{\partial L}{\partial \dot{q}} \right) \right) T'\end{aligned}$$

Recall that the Lagrangian can transform by a first derivative and preserve the action: $\delta \mathcal{L} = \frac{dg}{d\tau} = \frac{dg}{dt} T'$

$$\begin{aligned}\delta \mathcal{L} &= \frac{d}{dt} \left(\delta q \frac{\partial L}{\partial \dot{q}} + \delta t \left(L - \dot{q} \frac{\partial L}{\partial \dot{q}} \right) \right) T' = T' \frac{dg}{dt} \\ &= \frac{d}{dt} \left(\delta q \frac{\partial L}{\partial \dot{q}} + \delta t \left(L - \dot{q} \frac{\partial L}{\partial \dot{q}} \right) - g(q, t) \right) T' = 0\end{aligned}$$

Assuming a non-trivial T' , the time derivative must be zero, so the term inside must be constant:

$$\begin{aligned}\delta q \frac{\partial L}{\partial \dot{q}} + \delta t \left(L - \dot{q} \frac{\partial L}{\partial \dot{q}} \right) - g(q, t) &= \text{const.} \\ p\delta q - H\delta t - g(q, t) &= \text{const.}\end{aligned}\tag{10}$$

Therefore, if a time translation changes the Lagrangian by a time derivative, then it is a Noether symmetry with a corresponding conserved quantity $p\delta q - H\delta t - g(q, t)$.

3.2.4 Time Translational Symmetry and Energy Conservation

We shall consider a general Lagrangian $L(q, \dot{q}, t)$, under an infinitesimal time translation, the coordinates change by: $t \rightarrow t' = t + \epsilon$, $\delta t = \epsilon$. And position: $q \rightarrow q' = q$, $\delta q = 0$, $\delta \dot{q} = 0$. The change in the Lagrangian is:

$$\begin{aligned}\delta L &= \delta q \frac{\partial L}{\partial q} + \delta \dot{q} \frac{\partial L}{\partial \dot{q}} + \delta t \frac{\partial L}{\partial t} \rightarrow (0) \frac{\partial L}{\partial q} + (0) \frac{\partial L}{\partial \dot{q}} + \epsilon \frac{\partial L}{\partial t} \\ \delta L &= \epsilon \frac{\partial L}{\partial t}\end{aligned}$$

Recall that for time translation to be a symmetry, it must change the Lagrangian by at most a total time derivative $\delta L = \frac{dg}{dt}$. Equating the two, we obtain: $\epsilon \frac{\partial L}{\partial t} = \frac{dg}{dt}$. If the Lagrangian is independent of time, then $\frac{\partial L}{\partial t} = 0$ which means $g(q, t) = \text{const.}$ Applying these results into Eq. (10):

$$\begin{aligned}p\delta q - H\delta t - g &\rightarrow p(0) - \epsilon H - g = \text{const.} \\ H &= \text{const.}\end{aligned}$$

If time translation is a symmetry of the action, and the Lagrangian is time independent, then the Hamiltonian is conserved. We can identify the physical meaning of the Hamiltonian. Recall the definition of the Hamiltonian in terms of the canonical momentum $p_i = \frac{\partial L}{\partial \dot{q}_i}$:

$$H = \sum_i p_i \dot{q}_i - L.$$

Substituting the explicit form of the Lagrangian $L = T(\dot{q}) - V(q)$ into the definition of momentum:

$$p_i = \frac{\partial}{\partial \dot{q}_i} \left(\sum_j \frac{1}{2} m_j \dot{q}_j^2 - V(\vec{q}) \right) = \frac{\partial}{\partial \dot{q}_i} \left(\sum_j \frac{1}{2} m_j \dot{q}_j^2 \right) = m_i \dot{q}_i$$

Now, we substitute this momentum and the Lagrangian back into the expression for H :

$$\begin{aligned}H &= \sum_i (m_i \dot{q}_i) \dot{q}_i - \left(\sum_i \frac{1}{2} m_i \dot{q}_i^2 - V(\vec{q}) \right) \rightarrow \sum_i m_i \dot{q}_i^2 - \sum_i \frac{1}{2} m_i \dot{q}_i^2 + V(\vec{q}) \\ H &= \sum_i \frac{1}{2} m_i \dot{q}_i^2 + V(\vec{q})\end{aligned}$$

Recognizing that the first term is the Kinetic Energy (T) and the second is the Potential Energy (V), we conclude:

$$H = T + V = E_{\text{total}}.$$

Thus, for a system where the Lagrangian is time-independent, kinetic energy is quadratic with velocity, and potential is only position dependent, the conserved Noether charge resulting from time-translation symmetry is the Total Energy of the system. Helpful resources: [Dr. Rubens' Notes](#), and [Dr. Diaz's video](#).

4 Canonical Transformations

4.1 Derivation

In the Hamiltonian formulation, the position and momentum variables are independent of each other—they're on equal footing. This allows for more general coordinate transformations: $(q, p) \rightarrow (Q(q, p, t), P(q, p, t))$. Unlike a point transformation, the new coordinates are a mix of the old ones. This transformation is called "canonical" if the new Hamiltonian, $K(Q, P, t)$ preserves the form of Hamilton's equations:

$$\dot{q}_i = \frac{\partial H}{\partial p_i}, -\dot{p}_i = \frac{\partial H}{\partial q_i} \rightarrow \dot{Q}_i = \frac{\partial H}{\partial P_i}, -\dot{P}_i = \frac{\partial H}{\partial Q_i}$$

Recall from section 3 that the Lagrangian can change at most by a total time derivative, $\frac{dF}{dt}$, to preserve $\delta S = 0$ (Eq. (8)). This means a canonical transformation must also satisfy Hamilton's Principle ($\delta S' = 0$). The transformed Lagrangian is given by:

$$\begin{aligned} L' \equiv L'(Q, \dot{Q}, t) &= L(q, \dot{q}, t) - \frac{dF}{dt} \quad (- \text{ sign by convention}) \\ &= \sum P \dot{Q} - K(Q, P, t) = \sum p \dot{q} - H(q, p, t) - \frac{dF}{dt} \\ &\rightarrow \sum P dQ - K(Q, P, t) dt = \sum p dq - H(q, p, t) dt - dF \\ &\rightarrow dF = \sum p dq - \sum P dQ + (K - H) dt \end{aligned}$$

The differential dF is a linear combination of the above differentials, implying F must be of the form: $F \equiv F(q, Q, t)$. So its total differential must be:

$$dF(q, Q, t) = \sum \left(dq \frac{\partial F}{\partial q} + dQ \frac{\partial F}{\partial Q} \right) + \frac{\partial F}{\partial t}$$

Equating the differentials:

$$dF = \sum \left(dq \frac{\partial F}{\partial q} + dQ \frac{\partial F}{\partial Q} \right) + \frac{\partial F}{\partial t}$$

$$dF = \sum p dq - \sum P dQ + (K - H) dt \tag{11}$$

$$\rightarrow p = \frac{\partial F}{\partial q}, -P = \frac{\partial F}{\partial Q}, K = H + \frac{\partial F}{\partial t} \tag{12}$$

Where $F(q, Q, t)$ is called a type 1 generating function, denoted by $F_1(q, Q, t)$. There are 4 types of generating functions, but the most useful one is type 2: $F_2(q, P, t)$. We can obtain it by adding $d(\sum PQ)$ to dF :

$$\begin{aligned} d(\sum PQ) &= \sum (QdP + PdQ) \\ dF + d(\sum PQ) &= \sum p dq - \sum P dQ + (K - H) dt + \sum (QdP + PdQ) \\ &= \sum p dq + \sum QdP + (K - H) dt \equiv dF_2 \end{aligned}$$

The differential dF_2 is a linear combination of differentials dq, dP, dt so $F_2 \equiv F_2(q, P, t)$. Its total differential is given by:

$$dF_2 = \sum \left(dq \frac{\partial F_2}{\partial q} + dP \frac{\partial F_2}{\partial P} \right) + \frac{\partial F_2}{\partial t}$$

By equating differentials, the canonical transformations generated by F_2 are given by:

$$\boxed{p = \frac{\partial F_2}{\partial q}, Q = \frac{\partial F_2}{\partial P}, K = H + \frac{\partial F_2}{\partial t}} \quad (13)$$

4.2 Infinitesimal Canonical Transformations

Consider the infinitesimal dynamics of a Hamiltonian system, the evolution must be given by a canonical transformation because Hamilton's equations govern its dynamics, thus they're preserved for all time. The system should therefore evolve by:

$$Q = q + \delta q, P = p + \delta p$$

Because δq and δp are infinitesimal, the new variables only differ infinitesimally from the old. This implies the generating function that transforms the coordinates must be "near identity". Type 2 generating functions are perfect for this: Generating function: $F_2(q, P, t)$, canonical transformation: $p = \partial_q F_2, Q = \partial_P F_2$, Simplest transformation: $F_2 = qP \rightarrow Q = q, P = p$ (identity). We can construct a "near-identity" generating function as:

$$\begin{aligned} F &= F_2 + \epsilon G \\ \rightarrow F(q, P, t) &= qP + \epsilon G(q, P, t) \end{aligned} \quad (14)$$

Where G is the infinitesimal generating function. The transformations are:

$$\begin{aligned} p &= \partial_q(qP + \epsilon G) = P + \epsilon \partial_q G \rightarrow \delta p = \epsilon \partial_q G \\ q &= \partial_P(qP + \epsilon G) = q + \epsilon \partial_P G \rightarrow \delta q = \epsilon \partial_P G \\ \rightarrow \delta p &= \epsilon \partial_q G, \delta q = \epsilon \partial_P G \end{aligned} \quad (15)$$

Therefore, for an infinitesimal canonical transformation given by $F(q, P, t) = qP + \epsilon G(q, P, t)$ by Eq. (14), the canonical transformations are $p = P + \epsilon \partial_q G$, and $q = q + \epsilon \partial_P G$. The infinitesimal change in the coordinates are $\delta p = \epsilon \partial_q G$ and $\delta q = \epsilon \partial_P G$ as in Eq. (15).

4.3 The Hamiltonian and Infinitesimal Time Evolution

The evolution of coordinates $(q, p)|_{t=a}$ to $(Q, P)|_{t=b}$ must be canonical, as argued in section 4.2, and we've derived the form of the generating function for such infinitesimal transformations (Eqs. (14, 15)). For an infinitesimal canonical transformation through time, the system should evolve by:

$$q(t + dt) \approx q(t) + \dot{q}(t)dt, \quad p(t + dt) \approx p(t) + \dot{p}(t)dt$$

Where we applied a 1st order Taylor expansion centered at (q, p, t) to each coordinate. Because the system follows physical trajectories, it satisfies Hamilton's principle, and therefore is governed by Hamilton's equations. Substituting the expressions for \dot{q}, \dot{p} : $q(t + dt) \approx q(t) + \frac{\partial H}{\partial p} dt$, and $p(t + dt) \approx p(t) - \frac{\partial H}{\partial q} dt$. Thus yielding

$$\rightarrow \delta q = \frac{\partial H}{\partial p} dt, \quad \delta p = -\frac{\partial H}{\partial q} dt$$

Recall the general results for infinitesimal canonical transformations from Eqs.(14, 15), if we compare the results to the variations for the coordinates from time evolution then we see that $\epsilon = dt$ and:

$$p : \frac{\partial G}{\partial q} = \frac{\partial H}{\partial p}, \quad q : \frac{\partial G}{\partial P} = \frac{\partial H}{\partial p}$$

Proving the partials WRT time of G and H will prove they are equal up to a constant. Recall from Eq.(13) that $K = H + \frac{\partial F_2}{\partial t}$, using this we get:

$$F = qP + \epsilon G(q, P, t) \rightarrow K = H + dt \frac{\partial G}{\partial t}$$

For K to be an infinitesimal time evolved Hamiltonian, $K(Q, P, t) = H(Q, P, t + dt)$, we can Taylor expand $H(Q, P, t + dt)$ at (q, p, t) up to first order:

$$\begin{aligned} H(Q, P, t + dt) &\approx H(q, p, t) + (Q - q) \frac{\partial H}{\partial q} + (P - p) \frac{\partial H}{\partial p} + (t + dt - t) \frac{\partial H}{\partial t} \\ &= H(q, p, t) + \delta q \frac{\partial H}{\partial q} + \delta p \frac{\partial H}{\partial p} + dt \frac{\partial H}{\partial t} \end{aligned}$$

We found that $\delta q = \frac{\partial H}{\partial p} dt$, $\delta p = -\frac{\partial H}{\partial q} dt$, substituting:

$$\begin{aligned} H(Q, P, t + dt) &\approx H(q, p, t) + dt \frac{\partial H}{\partial p} \frac{\partial H}{\partial q} - dt \frac{\partial H}{\partial q} \frac{\partial H}{\partial p} + dt \frac{\partial H}{\partial t} \\ &= H(q, p, t) + dt \frac{\partial H}{\partial t} \end{aligned}$$

Equating the results: $K = H + dt \frac{\partial G}{\partial t}$ and $K(Q, P, t) = H(Q, P, t + dt) \approx H(q, p, t) + dt \frac{\partial H}{\partial t}$, we obtain:

$$\begin{aligned} H(q, p, t) + dt \frac{\partial G}{\partial t} &= H(q, p, t) + dt \frac{\partial H}{\partial t} \\ \rightarrow \frac{\partial G}{\partial t} &= \frac{\partial H}{\partial t} \end{aligned}$$

All partials of G and H are equal, thus proving that $G(q, P, t) = H(q, P, t) + \text{const.}$ So the generating function that transforms the system through time is given by $F_2(q, P, t) = qP + dt \frac{\partial H}{\partial t}$, which therefore proves the Hamiltonian is the generating function that infinitesimally evolves the system through time. This result should come as no surprise as it was proven in section 3, but here we prove it in the context of canonical transformations.

4.4 The Action and Time Evolution

By Hamilton's Principle, for any two points in configuration space, (q, t) and (Q, T) , there exists a unique physical trajectory—the "true path"—that satisfies the stationarity condition $\delta S = 0$. While we typically use this principle to find the specific path taken between fixed endpoints, we now shift our perspective to treat the Action as a function of the endpoints themselves. By considering incremental changes to both the starting point (q, t) and the destination (Q, T) , we create a more general "action field" where $S \equiv S(Q, T, q, t)$ is a scalar field defined by the physical path connecting these two events. Our objective is to analyze how the Action responds to these simultaneous variations by deriving its total differential dS , which will reveal how the initial and final momenta and energies are intrinsically linked to the geometry of this field. Fixing the times (t, T) , we consider the variation of S along a physical path where the end points vary:

$$\begin{aligned} S(q, t, Q, T) &= \int_{q, t}^{Q, T} L dt \rightarrow \delta S = [p(t) \delta q(t)]_t^T + \int_t^T dt \left(\frac{\partial L}{\partial q} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}} \right) \delta q \\ &= p(T) \delta q(T) - p(t) \delta q(t) \\ \delta S &= P \delta Q - p \delta q \end{aligned}$$

Where we have defined $p(T)\delta q(T) = P\delta Q$ and $p(t)\delta q(t) = p\delta q$ for convenience. We can equate this expression to the variation to obtain the expressions for the canonical transformations:

$$\begin{aligned}\delta S(q, t, Q, T) &= \delta q \frac{\partial S}{\partial q} + \delta t \frac{\partial S}{\partial t} + \delta Q \frac{\partial S}{\partial Q} + \delta T \frac{\partial S}{\partial T} \\ &= \delta q \frac{\partial S}{\partial q} + \delta Q \frac{\partial S}{\partial Q}\end{aligned}$$

Equating:

$$\begin{aligned}P\delta Q - p\delta q &= \delta q \frac{\partial S}{\partial q} + \delta Q \frac{\partial S}{\partial Q} \\ \rightarrow P &= \frac{\partial S}{\partial Q}, \quad p = -\frac{\partial S}{\partial q}\end{aligned}$$

Now consider varying the times (t, T) -the departure and arrival times. For $T \rightarrow T + dT$:

$$\begin{aligned}S(T) &= \int_t^{T+dT} L d\tau \rightarrow \text{FTC} \rightarrow \frac{dS}{dT} = \lim_{dT \rightarrow 0} \left(\frac{S(T+dT) - S(T)}{dT} \right) \\ &= \lim_{dT \rightarrow 0} \frac{1}{dT} \left(\int_t^{T+dT} L d\tau - \int_t^T L d\tau \right) \\ &= \lim_{dT \rightarrow 0} \frac{1}{dT} \left(\int_T^{T+dT} L d\tau \right) \\ &= L(q, \dot{q}, T) \frac{dT}{dT} = L(q, \dot{q}, T) \equiv L_T \\ \frac{dS}{dT} &= L_T\end{aligned}$$

Calculating the total derivative of $S(q, t, Q, T)$ WRT T :

$$\frac{dS(q, t, Q, T)}{dT} = \frac{dQ}{dT} \frac{\partial S}{\partial Q} + \frac{\partial S}{\partial T} \rightarrow P\dot{Q} + \frac{\partial S}{\partial T}$$

Equating:

$$\begin{aligned}L_T &= P\dot{Q} + \frac{\partial S}{\partial T} \rightarrow \frac{\partial S}{\partial T} = L_T - P\dot{Q} \\ \frac{\partial S}{\partial T} &= -H_T\end{aligned}$$

Performing the same procedure for $t \rightarrow t + dt$:

$$\begin{aligned}S(t) &= \int_{t+dt}^T L d\tau \rightarrow \text{FTC} \rightarrow \frac{dS}{dt} = \lim_{dt \rightarrow 0} \left(\frac{S(t+dt) - S(t)}{dt} \right) = \lim_{dt \rightarrow 0} \frac{1}{dt} \left(\int_{t+dt}^T L d\tau - \int_t^T L d\tau \right) \\ &= \lim_{dt \rightarrow 0} \frac{1}{dt} \left(\int_{t+dt}^T L d\tau - \left[\int_{t+dt}^T L d\tau + \int_t^{t+dt} L d\tau \right] \right) \\ &= \lim_{dt \rightarrow 0} \frac{1}{dt} \left(- \int_t^{t+dt} L d\tau \right) = -L(q, \dot{q}, t) \frac{dt}{dt} = -L(q, \dot{q}, T) \equiv -L_t \\ \frac{dS}{dt} &= -L_t\end{aligned}$$

Calculating the total derivative of $S(q, t, Q, T)$ WRT t :

$$\frac{dS(q, t, Q, T)}{dt} = \frac{dq}{dt} \frac{\partial S}{\partial q} + \frac{\partial S}{\partial t} \rightarrow -p\dot{q} + \frac{\partial S}{\partial t}$$

Equating:

$$\begin{aligned}-L_t &= -p\dot{q} + \frac{\partial S}{\partial t} \rightarrow \frac{\partial S}{\partial t} = p\dot{q} - L_t \\ \frac{\partial S}{\partial t} &= H_t\end{aligned}$$

Combining all the results to construct the total differential of S :

$$\begin{aligned} dS(q, t, Q, T) &= dq \frac{\partial S}{\partial q} + dt \frac{\partial S}{\partial t} + dQ \frac{\partial S}{\partial Q} + dT \frac{\partial S}{\partial T} \\ &= dq(-p) + dt(H_t) + dQ(P) + dT(-H_T) \end{aligned}$$

Because we are tracking the evolution of the system, the Lagrangians at the endpoints must evolve in time at the same rate through configuration space, this allows us to equate $dT = dt$ allowing for the following result:

$$dS = PdQ - pdq + (H_t - H_T)dt$$

For multiple degrees of freedom, dS becomes:

$$dS = \sum PdQ - \sum pdq + (H_t - H_T)dt \quad (16)$$

This is exactly the differential for the F_1 generating function in Eq.(11): $dF = \sum pdq - \sum PdQ + (K - H)dt$, off by a (-). We conclude that $-dS = dF_1$, which implies that the action S is actually the generating function that canonically transforms the Hamiltonian through time, specifically $-S$ generates forward time evolution.

5 Hamilton-Jacobi Theory

5.1 Derivation

In the previous section, we transformed the Action from a static integral into a dynamic scalar field $S(q, t, Q, T)$ by treating the initial and final position and the departure and arrival times as independent variables. The key point is that the physical evolution of a system through time is actually a continuous canonical transformation, with the Action S serving as the unique generating function that "unfolds" the initial conditions into the future state. This provides us with a new lens to view the dynamics of mechanical systems—through the Hamilton-Jacobi equation. We begin with the action functional $S[q(t)]$, which returns a scalar given any path $q(t)$ —physical or not. Given a the initial conditions, (q_0, t_0) , we can calculate the action of all physical paths by fixing (q_0, t_0) and varying the final points (q, t) . We can represent this with "Hamilton's Principle Function":

$$S(q, t) = \int_{q_0, t_0}^{q, t} L dt'$$

Note The key difference: $S[q(t)]$ is a functional— takes any path $q(t)$ for an input and $S(q, t)$ is a function with that calculates the action for physical paths only, given .

Similar to the previous section, if we fix the variation in time, and vary the endpoints, we can see how the action responds to such variations;

$$\begin{aligned} \delta S(q, t) &= \delta \int_{q_0, t_0}^{q, t} L dt' = \int_{t_0}^t \left(\delta q \frac{\partial L}{\partial q} + \delta \dot{q} \frac{\partial L}{\partial \dot{q}} + 0 \right) dt' \\ &= \int_{t_0}^t \left(\delta q \frac{d}{dt'} \frac{\partial L}{\partial \dot{q}} + \delta \dot{q} \frac{\partial L}{\partial \dot{q}} \right) dt' = \int_{t_0}^t \frac{d}{dt'} \left(\delta q \frac{\partial L}{\partial \dot{q}} \right) dt' \\ &= \int_{t_0}^t \frac{d}{dt'} (p(t') \delta q) dt' = [p(t') \delta q(t')]_{t_0}^t \\ &= p(t) \delta q(t) \end{aligned}$$

We can equate this to the definition of the variation of S : $\delta S = \delta q \frac{\partial S}{\partial q} + \delta t \frac{\partial S}{\partial t} = \delta q \frac{\partial S}{\partial q} + 0$

$$\delta q \frac{\partial S}{\partial q} = p(t) \delta q(t) \rightarrow p(t) = \frac{\partial S}{\partial q}$$

To find the explicit time dependence of S , we examine its total differential: $dS(q, t) = dq \frac{\partial S}{\partial q} + dt \frac{\partial S}{\partial t} = pdq + dt \frac{\partial S}{\partial t}$ with:

$$\begin{aligned} S &= \int L dt \rightarrow dS = L dt = pdq - H dt \\ pdq - H dt &= pdq + dt \frac{\partial S}{\partial t} \\ \rightarrow \frac{\partial S}{\partial t} &= -H \end{aligned}$$

We obtained the following $p(t) = \frac{\partial S}{\partial q}$, $-H(q, p, t) = \frac{\partial S}{\partial t}$, we can further simplify this into one result by substituting the momentum into H which nets us:

$$\boxed{H\left(q, \frac{\partial S}{\partial q}, t\right) + \frac{\partial S}{\partial t} = 0} \quad (17)$$

This is the Hamilton-Jacobi Equation—for all physical paths the time dependence of the action is governed by the Hamiltonian. Since this is a first order PDE of n coordinates and time, the full solution is the complete integral which is a function of the n independent variables, and n arbitrary constants for each variable. For the action, we have s number of coordinates and 1 time—a total of $s + 1$ variables which need $s + 1$ constants. Letting α_i represent these constants, the complete integral that solves the HJE (Eq. (17)) is given by $S = F(q, \alpha, t) + A$ where A is a constant, and the subscripts were dropped for conciseness.

To extract the useful dynamics from this formulation, if we apply a canonical transformation from coordinates (q, p) to (Q, P) using $F(q, \alpha, t)$ as the generating function, and choosing to treat the α as the new momentum (or equivalently as the new position). Since the generating function depends on old position and new momentum, we choose an $F_2(q, P, t)$, with canonical transformations

$$\begin{aligned} p &= \partial_q F_2 \rightarrow p + \partial_q A = \partial_q (F + A) \rightarrow p = \partial_q S \\ Q &= \partial_P F_2 \rightarrow Q = \partial_P S \\ \rightarrow p &= \partial_q S, \quad Q = \partial_P S \end{aligned} \quad (18)$$

and the transformed Hamiltonian: $K = H + \partial_t(F) \rightarrow K + \partial_t A = H + \partial_t F + \partial_t A \rightarrow K = H + \partial_t S \rightarrow K = 0$. This means that the transformed Hamiltonian K must be zero. Applying the canonical equations: $-\dot{\alpha} = \frac{\partial K}{\partial \dot{\alpha}}, \dot{Q} = \frac{\partial K}{\partial Q} \rightarrow \dot{\alpha} = 0, \dot{Q} = 0$, thus:

$$\alpha = \text{const.} \quad Q = \text{const.} \quad (19)$$

Effectively, the canonical transformations are given by: $p = \partial_q S$, $Q = \partial_P S$, and $K = 0$. Showing this explicitly reconnects to the prior results in section 4 where we show the action is the generating function that canonically transforms the coordinates through time, here we re derive the result in another context.

To summarize, we constructed $S(q, t)$ which gives the action for all physical paths and that they all must obey the HJE. We found that solving the HJE gives $S(q, P, t)$ or $S(q, Q, t)$, and that S is the generating function that canonically transforms the coordinates such that the new Hamiltonian is zero ($K = 0$), making the motion of the new coordinates stationary ($\dot{\alpha} = 0, \dot{Q} = 0$). This means that if we solve the HJE for S , we can calculate the dynamics of the old coordinates using the canonical transformations $p = \partial_q S$, and $Q = \partial_P S$.

5.2 Time Independent HJE (TIHJE)

We have previously established that if the Hamiltonian is independent of time, and quadratic with momentum, then it represents the total energy of the system. Using the separability ansatz that

$$S(q, \alpha, t) = W(q, \alpha) - \alpha t \quad (20)$$

where W is referred to as "Hamilton's Characteristic Function", defined as

$$W(q_i, \alpha_i) = \int \sum_i p_i dq_i$$

with differential

$$dW(q_i, \alpha_i) = \sum_i \frac{\partial W}{\partial q_i} dq_i$$

where the $d\alpha_i = 0$ because they are constant. The HJE becomes:

$$\begin{aligned} H\left(q, \frac{\partial S}{\partial q}, t\right) + \frac{\partial S}{\partial t} = 0 &\rightarrow H\left(q, \frac{\partial W}{\partial q}\right) + \frac{\partial(-\alpha_t t)}{\partial t} = 0 \rightarrow H = \alpha_t \equiv E \\ &\rightarrow H\left(q, \frac{\partial W}{\partial q}\right) = E \end{aligned} \quad (21)$$

The method is to now solve the TIHJE (Eq. 21) but this time using $W \equiv W(q, \alpha)$ as a type 2 generating function with transformations: $p_i = \partial_{q_i} W, Q_i = \partial_{P_i} W$. The transformed Hamiltonian will be non zero:

$$\begin{aligned} K = H + \frac{\partial W(q, \alpha)}{\partial t} &\rightarrow K = H + 0 = E \\ K &= H = E \end{aligned}$$

With a constant new Hamiltonian, the transformed momenta are:

$$\begin{aligned} \dot{P} &= -\frac{\partial K}{\partial Q_i} = -\frac{\partial(E)}{\partial Q_i} = 0 \\ P_i &= \text{const.} \equiv \alpha_i \end{aligned}$$

However the new positions Q_i may no longer constant but linear with time:

$$\dot{Q}_i = \frac{\partial K}{\partial \alpha_i} = \frac{\partial(\alpha_j)}{\partial \alpha_i} = \begin{cases} 1, i = j \\ 0, i \neq j \end{cases} \rightarrow \begin{cases} Q_i = \beta_i + t = \frac{\partial W}{\partial E}, i = j \\ Q_i = \beta_i = \frac{\partial W}{\partial \alpha_i}, i \neq j \end{cases}$$

5.2.1 Example with Harmonic Oscillator

For the Harmonic oscillator in Cartesian coordinates, the Hamiltonian is independent of time (thus representing the total energy) and is given by $H(q, p) = \frac{1}{2m}(p^2 + (mwq)^2) = E$. Using the canonical transformations for the characteristic function $p_i = \partial_{q_i} W, Q_i = \partial_{P_i} W$, the HJE becomes:

$$\begin{aligned} H\left(q, \frac{\partial W}{\partial q}\right) &= E \\ \frac{1}{2m} \left[\left(\frac{\partial W}{\partial q} \right)^2 + (mwq)^2 \right] &= E \end{aligned}$$

Because there is only one coordinate, the partial becomes a total derivative:

$$\begin{aligned} \left(\frac{dW}{dq} \right)^2 &= 2mE - (mwq)^2 \\ W &= \int \sqrt{2mE - (mwq)^2} dq \end{aligned}$$

The new position coordinate (there is only one) is given by $Q = \beta + t = \frac{dW}{d\alpha_t}$, where $\alpha_t = E$;

$$\begin{aligned} \beta + t &= \frac{\partial W}{\partial \alpha_t} = \frac{\partial}{\partial \alpha_t} \left(\int \sqrt{2mE - (mwq)^2} dq \right) = \frac{1}{w} \arcsin \left(q \sqrt{\frac{mw^2}{2E}} \right) \\ \rightarrow q(t) &= \sqrt{\frac{2E}{mw^2}} \sin(wt + \phi) \end{aligned} \quad (22)$$

where $\phi = w\beta$. From here the old momentum is found using the expression for q and the canonical transformation:

$$\begin{aligned}
p &= \frac{\partial}{\partial q} \left(\int \sqrt{2mE - m^2 w^2 q^2} dq \right) \\
&= \sqrt{2mE - m^2 w^2 q^2} \\
&= \sqrt{2mE - m^2 w^2 \left[\sqrt{\frac{2E}{mw^2}} \sin(wt + \phi) \right]^2} \\
&\rightarrow p(t) = \sqrt{2mE} \cos(wt + \phi)
\end{aligned} \tag{23}$$

5.3 Application to The Kepler Problem

We consider a mass m moving under the influence of a central gravitational potential $V(r)$, like the Earth-Moon system for example. We want to extract the dynamics of such a system, where the larger mass is infinite so its stationary and the other mass orbits. The goal is to obtain the orbital equation that describes the path in space, and Kepler's equation which describes the position through time. The Hamilton for this system is:

$$H(q, p) = \frac{1}{2m} \left(p_r^2 + \frac{p_\theta^2}{r^2} + \frac{p_\phi^2}{r^2 \sin^2 \theta} \right) - \frac{k}{r}$$

Because the Hamiltonian is independent of time, $H(q, p) = E$, so we can apply TIHJE to solve the dynamics:

$$\begin{aligned}
H \left(q, \frac{\partial W}{\partial q} \right) &= E \\
\frac{1}{2m} \left(\left(\frac{\partial W}{\partial r} \right)^2 + \frac{1}{r^2} \left(\frac{\partial W}{\partial \theta} \right)^2 + \frac{1}{r^2 \sin^2 \theta} \left(\frac{\partial W}{\partial \phi} \right)^2 \right) - \frac{k}{r} &= E
\end{aligned}$$

We apply the separation of variables technique and assume: $W = W_r(r) + W_\theta(\theta) + W_\phi(\phi)$. As a result, the partial derivatives $\frac{\partial W}{\partial q_i}$ become total derivatives:

$$\frac{1}{2m} \left(\left(\frac{dW_r}{dr} \right)^2 + \frac{1}{r^2} \left(\frac{dW_\theta}{d\theta} \right)^2 + \frac{1}{r^2 \sin^2 \theta} \left(\frac{dW_\phi}{d\phi} \right)^2 \right) - \frac{k}{r} = E$$

Solving for each W_i , we start with the ϕ coordinate, and apply $\frac{\partial}{\partial \phi}()$ to the HJE:

$$\begin{aligned}
\frac{\partial}{\partial \phi}(H) &= \frac{\partial}{\partial \phi}(E) \\
\frac{1}{r^2 \sin^2 \theta} \frac{\partial}{\partial \phi} \left[\left(\frac{dW_\phi}{d\phi} \right)^2 \right] &= 0 \rightarrow \frac{\partial}{\partial \phi} \left[\left(\frac{dW_\phi}{d\phi} \right)^2 \right] = 0 \\
\left(\frac{dW_\phi}{d\phi} \right)^2 &= \text{const.} \rightarrow \frac{dW_\phi}{d\phi} = \alpha_\phi \\
\boxed{W_\phi &\equiv \alpha_\phi \phi}
\end{aligned}$$

We substitute $\frac{dW_\phi}{d\phi} = \alpha_\phi$ and apply $\frac{\partial}{\partial \theta}()$ to the HJE to solve for W_θ : $\frac{\partial}{\partial \theta}(H) = \frac{\partial}{\partial \theta}(E) \rightarrow$

$$\begin{aligned}
\frac{1}{r^2} \frac{\partial}{\partial \theta} \left(\left(\frac{dW_\theta}{d\theta} \right)^2 + \frac{\alpha_\phi^2}{\sin^2 \theta} \right) &= 0 \\
\left(\frac{dW_\theta}{d\theta} \right)^2 + \frac{\alpha_\phi^2}{\sin^2 \theta} &= \text{const.} \equiv \alpha_\theta^2 \\
\boxed{W_\theta = \int \sqrt{\alpha_\theta^2 - \frac{\alpha_\phi^2}{\sin^2 \theta}} d\theta}
\end{aligned} \tag{24}$$

Lastly for W_r , we substitute $\left(\frac{dW_\theta}{d\theta}\right)^2 + \frac{\alpha_\phi^2}{\sin^2\theta} = \alpha_\theta^2$ to the HJE, and we don't need to differentiate, the HJE is already function of r with these substitutions:

$$\frac{1}{2m} \left[\left(\frac{dW_r}{dr} \right)^2 + \frac{\alpha_\theta^2}{r^2} \right] - \frac{k}{r} = E$$

$$\boxed{W_r = \int \sqrt{2mE + 2mk/r - \alpha_\theta^2/r^2} dr} \quad (25)$$

So the complete integral (generating function) is given by:

$$W = \int \sqrt{2mE + 2mk/r - \alpha_\theta^2/r^2} dr + \int \sqrt{\alpha_\theta^2 - \frac{\alpha_\phi^2}{\sin^2\theta}} d\theta + \alpha_\phi \phi$$

Now we can use the canonical transformations $Q_i = \beta_i = \frac{\partial W}{\partial \alpha_i}$ to extract the dynamics. Starting with the θ coordinate, we can obtain the orbital equation $r(\phi)$:

$$\begin{aligned} \beta_\theta &= \frac{\partial W}{\partial \alpha_\theta} = \frac{\partial}{\partial \alpha_\theta} \left(\int \sqrt{2mE + 2mk/r - \alpha_\theta^2/r^2} dr + \int \sqrt{\alpha_\theta^2 - \frac{\alpha_\phi^2}{\sin^2\theta}} d\theta \right) + 0 \\ &= \int \frac{\alpha_\theta d\theta}{\sqrt{\alpha_\theta^2 - \alpha_\phi^2/\sin^2\theta}} + \int \frac{(-\alpha_\theta/r^2) dr}{\sqrt{2mE + 2mk/r - \alpha_\theta^2/r^2}} \end{aligned}$$

To simplify, we orient the coordinates so the orbit is in the equatorial plane ($\theta = \pi/2 \rightarrow \alpha_\theta = \alpha_\phi$). the 1st integral becomes ϕ , and the rest is:

$$\phi - \beta_\theta = \int \frac{(\alpha_\theta/r^2) dr}{\sqrt{2mE + 2mk/r - \alpha_\theta^2/r^2}} \rightarrow \frac{1}{r} = \frac{mk}{\alpha_\theta^2} (1 + \epsilon \cos(\phi - \beta_\theta))$$

with eccentricity $\epsilon = \sqrt{1 + \frac{2E\alpha_\theta^2}{mk^2}}$. With the last bit of algebra, we arrive at the familiar orbital equation:

$$\boxed{r = \frac{\alpha_\theta^2}{mk(1 + \epsilon \cos(\phi - \beta_\theta))}} \quad (26)$$

All that's left is to derive Kepler's equation, this is done by using the transformation for β_E :

$$\beta_E + t = \frac{\partial W}{\partial E} = \int \frac{m dr}{\sqrt{2mE + 2mk/r - \alpha_\theta^2/r^2}}$$

Here we must make the assumption that the energy $E < 0$ so the orbit is elliptical, then apply a substitution $r = a(1 - \epsilon \cos u)$, where $a = -k/2E$. The integral becomes:

$$\begin{aligned} \beta_E + t &= \sqrt{\frac{ma^3}{k}} \int (1 - \epsilon \cos u) du \\ &= \sqrt{\frac{ma^3}{k}} (u - \epsilon \sin u) \end{aligned}$$

We choose β_E so that at $t = 0$ the r coordinate is at a minimum, which implies u must be zero, and hence $\beta_E = 0$. We lastly define $n = \sqrt{k/ma^3}$ and finally arrive at

$$\boxed{nt = u - \epsilon \sin u} \quad (27)$$

Kepler's Equation.

6 Action-Angle Variables

In the standard Hamilton-Jacobi method, one typically seeks a canonical transformation that reduces the new Hamiltonian to zero, thereby "freezing" the motion into a set of constant coordinates. However, this approach is often limited by the mathematical complexity of the trajectories; in orbital mechanics, for instance, the resulting elliptic integrals frequently lack the analytical inversions necessary to solve for $q(t)$ explicitly.

For systems exhibiting periodic motion, where the phase-space trajectory traces a closed orbit, we can bypass these complexities by focusing on the underlying geometry rather than the explicit trajectory. This is achieved through the introduction of Action-Angle variables, (θ, J) , which act as a "polar" representation of the phase-space orbit. In this framework, we specifically require the new Hamiltonian to depend only on the momentum/action, $H(J)$. By making the Hamiltonian independent of the angle θ , we define θ as a cyclic coordinate, forcing it to evolve linearly in time. This setup aligns perfectly with physical intuition: the canonical momentum J serves as a constant "label" for the orbit's size and energy, while the canonical position θ acts as a phase variable that tracks the system's progress through its cycle.

This technique is uniquely powerful for two primary reasons. First, it allows us to calculate a system's fundamental frequencies directly from the Hamiltonian without ever solving the equations of motion for $q(t)$. Second, the action variable J is an adiabatic invariant, meaning it remains nearly constant even when the system's parameters are slowly varied. Consequently, Action-Angle variables provide the natural language for describing complex periodic systems, perturbations, and the historical transition from classical mechanics into the quantum regime (Bohr-Sommerfeld Quantization condition).

6.1 Derivation

Assuming the TIHJE: $H(q, p) = E$, we seek a canonical transformation to new variables (θ, J) such that the new Hamiltonian K depends only on the canonical momenta J , $K \equiv K(J)$. This means

$$-\dot{J} = \frac{\partial K}{\partial \theta} = 0 \rightarrow J = \text{const.}$$

We can use the characteristic function $W(q, J)$ as a generating function to accomplish this. The Hamiltonian transforms as

$$\begin{aligned} K(J) &= H + \partial_t W(q, J) \rightarrow K(J) = H = E \\ K(J) &= E(J) \end{aligned} \tag{28}$$

Here we show an important result that if the transformed Hamiltonian is only a function of the momenta J , then so must the energy (assuming time independent H). The coordinates transform as:

$$p = \frac{\partial W}{\partial q}, \quad \theta = \frac{\partial W}{\partial J} \tag{29}$$

Since we require $H(q, p) = E \rightarrow H(q, p) = K(J)$ we substitute the transformation for p to obtain $H(q, \frac{\partial W}{\partial q}) = K(J)$.

Separately, we can use $p = \frac{\partial W}{\partial q}$ to solve for the W by integrating its differential: $dW(q, J) = \frac{\partial W}{\partial q} dq + \frac{\partial W}{\partial J} dJ = \frac{\partial W}{\partial q} dq$ (J is constant). The integral is given by:

$$W(q, J) = \int_0^q p(q', J) dq'$$

Using the transformation equation $\theta = \frac{\partial W}{\partial J}$, we obtain an expression for θ :

$$\theta = \frac{\partial}{\partial J} \left(\int_0^q p(q', J) dq' \right)$$

At this stage, we don't know what J is, just that it's a constant. And we don't know the range of θ . We can find J by enforcing the idea that for periodic motion, the canonical position should increase by 2π

over a full cycle, which is expressed with $\Delta\theta = \oint d\theta = 2\pi$. We know that because $\theta = \frac{\partial W}{\partial J}$, $\theta \equiv \theta(q, J)$. So its differential is $d\theta = \frac{\partial\theta}{\partial q}dq + \frac{\partial\theta}{\partial J}dJ$, but again $J = \text{const.}$ so $d\theta = \frac{\partial\theta}{\partial q}dq$. Substituting this into $\Delta\theta$:

$$\begin{aligned}\Delta\theta &= \oint d\theta \rightarrow \oint \frac{\partial\theta}{\partial q}dq \rightarrow \oint \frac{\partial}{\partial q} \left(\frac{\partial W}{\partial J} \right) dq \\ &= \oint \frac{\partial}{\partial J} \left(\frac{\partial W}{\partial q} \right) dq \rightarrow \oint \frac{\partial}{\partial J} (p) dq\end{aligned}$$

Where we applied the definition of θ , swapped the order of differentiation (assuming smooth 2nd derivatives) and applied the definition of p . Next we apply Leibniz Rule (from appendix A) to obtain:

$$\Delta\theta = \oint \frac{\partial}{\partial J} (p) dq = \frac{d}{dJ} \oint p dq$$

Now we can enforce the periodicity condition that $\Delta\theta = \oint d\theta = 2\pi$:

$$\begin{aligned}\Delta\theta &= \frac{d}{dJ} \oint p dq = 2\pi \\ \int 2\pi dJ &= \int \left(\frac{d}{dJ} \oint p dq \right) dJ \\ J &= \frac{1}{2\pi} \oint p dq\end{aligned}$$

By requiring the canonical position change by 2π every cycle, we derived that the constant conjugate momenta, J , must be the area enclosed by the orbit in the phase space.

6.1.1 Frequency of Oscillation

Now we can move to extracting the usefulness of this method by obtaining the frequency of oscillation with just the canonical equations: $-\dot{J} = \frac{\partial K}{\partial \theta} = 0$, and $\dot{\theta} = \frac{\partial K}{\partial J} = \frac{\partial E(J)}{\partial J} \rightarrow \dot{\theta} = \frac{dE}{dJ}$. Note that E is a constant of motion, but its value depends on the momentum J , so $\frac{dE}{dJ} \neq 0$. We can show this is the frequency $\nu = 1/T$. Recall that $T = \oint dq/\dot{q} = \oint dq/(\partial H/\partial p)$. We will first calculate $\frac{dJ}{dE}$ by Leibniz Rule:

$$\frac{dJ}{dE} = \frac{d}{dE} \left(\frac{1}{2\pi} \oint p dq \right) = \frac{1}{2\pi} \oint \frac{\partial p}{\partial E} dq \quad (30)$$

Next we can calculate $\frac{\partial p}{\partial E}$ from the Hamiltonian: $H(q, p) = E$:

$$\begin{aligned}H(q, p) = E &\rightarrow \frac{\partial H}{\partial E} = \frac{\partial E}{\partial E} \\ \frac{\partial p}{\partial E} \frac{\partial H}{\partial p} &= 1 \rightarrow \frac{\partial p}{\partial E} = \frac{1}{\frac{\partial H}{\partial p}} \\ \frac{\partial p}{\partial E} &= \frac{1}{\dot{q}}\end{aligned}$$

now we can substitute this result into the integral:

$$\begin{aligned}\frac{dJ}{dE} &= \frac{1}{2\pi} \oint \frac{\partial p}{\partial E} dq = \frac{1}{2\pi} \oint \frac{1}{\dot{q}} dq \rightarrow \frac{1}{2\pi} \oint dt \\ &= \frac{1}{2\pi} \int_0^T dt = \frac{1}{2\pi} (T) \\ &= \frac{T}{2\pi}\end{aligned}$$

We can then use $w = 2\pi/T$ so that $\frac{dJ}{dE} = 1/w$. Lastly, we can use this result for the canonical position transformation:

$$\dot{\theta} = \frac{dE}{dJ} = w$$

Thus, the Action-Angle method of transforming to a momentum dependent Hamiltonian and enforcing periodicity on the canonical position allows us to bypass ever solving for the trajectories and only calculate the momenta J and differentiate to find the frequency of oscillation.

6.1.2 Harmonic Oscillator Revisited

Consider the Hamiltonian for a Simple Harmonic Oscillator:

$$H(q, p) = \frac{p^2}{2m} + \frac{1}{2}m\omega_0^2 q^2 = E$$

The turning points occur when $p = 0$, giving $q_{max} = \pm\sqrt{2E/m\omega_0^2}$. The action J is defined as the area of the orbit in phase space divided by 2π :

$$J = \frac{1}{2\pi} \oint p dq = \frac{(2)}{2\pi} \int_{-q_{max}}^{q_{max}} \sqrt{2mE - m^2\omega_0^2 q^2} dq$$

Where the factor of 2 accounts for the fact that the motion starts at $\pi/2 \rightarrow -\pi/2$ and then back to $\pi/2$. Using the substitution $q = \sqrt{\frac{2E}{m\omega_0^2}} \sin \phi$, where $dq = \sqrt{\frac{2E}{m\omega_0^2}} \cos \phi d\phi$, the limits transform from $[-q_{max}, q_{max}]$ to $[-\pi/2, \pi/2]$:

$$\begin{aligned} J &= \frac{2}{2\pi} \int_{-\pi/2}^{\pi/2} \sqrt{2mE(1 - \sin^2 \phi)} \left(\sqrt{\frac{2E}{m\omega_0^2}} \cos \phi \right) d\phi \\ J &= \frac{1}{\pi} \frac{2E}{\omega_0} \int_{-\pi/2}^{\pi/2} \cos^2 \phi d\phi = \frac{2E}{\pi\omega_0} \left[\frac{\phi}{2} + \frac{\sin 2\phi}{4} \right]_{-\pi/2}^{\pi/2} \\ J &= \frac{2E}{\pi\omega_0} \left(\frac{\pi}{2} \right) = \frac{E}{\omega_0} \end{aligned}$$

Solving for energy, the new Hamiltonian is $K(J) = \omega_0 J$. Hamilton's equations in the new variables (θ, J) are:

- $\dot{J} = -\frac{\partial K}{\partial \theta} = 0 \implies \mathbf{J} = \text{constant}$
- $\dot{\theta} = \frac{\partial K}{\partial J} = \omega_0 \implies \theta(\mathbf{t}) = \omega_0 \mathbf{t} + \beta$

Thus, J represents the conserved scaling of the orbit (proportional to Area), and θ is the linear phase of the oscillation.

6.1.3 Quartic Oscillator Example

To showcase the many uses of Action-Angle variables, lets examine the time-independent quartic oscillator, given by

$$H(q, p) = \frac{1}{2m}p^2 + \alpha q^4$$

and we want to find the frequency of oscillation. We know the period is given by $T = \oint dq/\dot{q} = \oint dq/(\partial H/\partial p) = 1/w$. So we begin by calculating \dot{q} using Hamilton's equations :

$$\dot{q} = (\partial H/\partial p) = p/m$$

So the period integral is given by: $T = \oint \frac{mdq}{p}$, but we still need an expression for p . We can use the fact that $H(q, p) = E$ to solve for p :

$$H(q, p) = \frac{1}{2m}p^2 + \alpha q^4 = E \rightarrow p = \sqrt{2m(E - \alpha q^4)}$$

Now substituting this into the period integral:

$$T = \oint \frac{mdq}{p} = \oint \frac{mdq}{\sqrt{2m(E - \alpha q^4)}}$$

and we are immediately stuck! This is an elliptic integral of the first kind with no analytic solutions :((((.

Action-angle variables offer insights into frequency analysis despite the elliptic integrals, we start by calculating the momenta/action variable

$$J = \frac{1}{2\pi} \oint p dq = \frac{1}{2\pi} \oint \sqrt{2m(E - \alpha q^4)} dq$$

This is still an elliptic integral, but we can extract some information about the oscillation of this system, lets apply a substitution given by $u = (\alpha/E)^{1/4} q \rightarrow du = (\alpha/E)^{1/4} dq$, then the momenta becomes:

$$J = \frac{\sqrt{2mE}}{2\pi} \left(\frac{E}{\alpha} \right)^{1/4} \oint \sqrt{1 - u^4} du$$

The substitution allowed us to examine how J scales with E as the integral is independent of E , so it must evaluate to some constant $A \equiv \oint \sqrt{1 - u^4} du$. The expression then becomes:

$$J = \frac{\sqrt{2mE}}{2\pi} \left(\frac{E}{\alpha} \right)^{1/4} A = CE^{3/4}$$

So $J = CE^{3/4}$. The frequency is given by $\dot{\theta} = \frac{dE}{dJ} = w$, but all we know is that $J \propto E^{3/4} \rightarrow E \propto J^{4/3}$. So then

$$\dot{\theta} = \frac{dE}{dJ} \propto \frac{d}{dJ}(J^{4/3}) \propto E^{1/4}$$

Which ultimately implies $w \propto E^{1/4}$. We may not have the exact relation but we know the behavior of the frequency; its dependence on energy is to the 1/4th power.

6.1.4 Kepler Orbits Revisited

The Kepler problem (a mass m in a potential $V = -k/r$) is the classic demonstration of how Action-Angle variables reveal hidden symmetries, such as why planetary orbits are closed ellipses. In spherical coordinates (r, θ, ϕ) , the Hamiltonian is:

$$H = \frac{1}{2m} \left(p_r^2 + \frac{p_\theta^2}{r^2} + \frac{p_\phi^2}{r^2 \sin^2 \theta} \right) - \frac{k}{r} = E$$

The system possesses three constants of motion: energy E , total angular momentum L , and the z -component L_z . These allow us to separate the Action integrals. We define three action variables $J_i = \frac{1}{2\pi} \oint p_i dq_i$:

- **Azimuthal Action:** Since $p_\phi = L_z$ is constant:

$$J_\phi = \frac{1}{2\pi} \int_0^{2\pi} L_z d\phi = L_z$$

- **Latitudinal Action:** Solving $L^2 = p_\theta^2 + p_\phi^2 / \sin^2 \theta$ for p_θ :

$$J_\theta = \frac{1}{2\pi} \oint \sqrt{L^2 - \frac{J_\phi^2}{\sin^2 \theta}} d\theta = L - J_\phi \implies L = J_\theta + J_\phi$$

- **Radial Action:** Substituting the total angular momentum L into the energy equation:

$$J_r = \frac{1}{2\pi} \oint \sqrt{2mE + \frac{2mk}{r} - \frac{(J_\theta + J_\phi)^2}{r^2}} dr$$

Using techniques from complex analysis to evaluate this complex integral over a full radial cycle:

$$J_r = \frac{mk}{\sqrt{-2mE}} - (J_\theta + J_\phi)$$

Solving for the energy E in terms of the actions, we find the new Hamiltonian $K(J)$:

$$K(J_r, J_\theta, J_\phi) = -\frac{mk^2}{2(J_r + J_\theta + J_\phi)^2}$$

Defining the principal action as $J_n = J_r + J_\theta + J_\phi$, we find the fundamental frequencies $\nu_i = \partial K / \partial J_i$:

$$\nu_r = \nu_\theta = \nu_\phi = \frac{mk^2}{(J_r + J_\theta + J_\phi)^3} = \nu$$

Because the Hamiltonian depends only on the sum of the actions, all three orbital frequencies are identical. This mathematical degeneracy is the fundamental reason planetary orbits are closed ellipses rather than complex precessing shapes. Furthermore, this result provided the basis for the Bohr-Sommerfeld quantization of the hydrogen atom, where $J_n = n\hbar$ yields the Rydberg formula.

6.2 Canonical Perturbation Theory

In real life we rarely encounter scenarios where the equations of motion are exactly solvable. But in many instances, a system is very close to an exactly solvable system, off by some disturbance or perturbation. In this section we derive the application of perturbative techniques to periodic systems. Given a Hamiltonian already expressed in action-angle variables (θ, J) , the unperturbed part is given by H_0 , and the perturbation is given by ϵH_1 , where ϵ is a small dimensionless parameter to control the strength of the perturbation. The total Hamiltonian is therefore given by:

$$H(\theta, J, t) = H_0(J) + \epsilon H_1(\theta, J, t)$$

For the application of action-angle variables, we will assume the perturbation is periodic. Just like the standard action-angle variables technique, we aim to transform total Hamiltonian $H(\theta, J, t)$ to new canonical coordinates (θ', J') so that the new Hamiltonian is independent of the canonical position θ' $K \equiv K(J', t)$. Then from the canonical equations it follows that $-\dot{J}' = \frac{\partial K(J', t)}{\partial \theta'} = 0 \rightarrow J' = \text{const.}$ We use a type 2 generating function $S(\theta, J', t)$ and expand this and the new Hamiltonian in powers of ϵ :

$$\begin{aligned} S(\theta, J', t) &= \theta J' + \epsilon S_1(\theta, J', t) + \dots \\ K(J', t) &= K_0(J') + \epsilon K_1(J', t) + \dots \end{aligned}$$

With canonical transformations:

$$\begin{aligned} J &= \frac{\partial S}{\partial \theta} = J' + \epsilon \frac{\partial S_1}{\partial \theta} + \dots \\ \theta' &= \frac{\partial S}{\partial J'} = \theta + \epsilon \frac{\partial S}{\partial J'} + \dots \\ K &= H(\theta, J, t) + \frac{\partial S}{\partial t} \end{aligned}$$

From the series expansion, we can use them to obtain the first order correction to desired quantities like the Hamiltonian K and the shifted frequencies w' . We need the corrected Hamiltonian first, by using the 1st order corrections to the following:

$$\begin{aligned} J &\approx J' + \epsilon \frac{\partial S_1}{\partial \theta} \\ \theta' &\approx \theta + \epsilon \frac{\partial S}{\partial J'} \\ S(\theta, J', t) &\approx \theta J' + \epsilon S_1(\theta, J', t) \\ K(J', t) &\approx K_0(J') + \epsilon K_1(J', t) \end{aligned}$$

Equating the canonical transformation for the Hamiltonian to its first order approximation we obtain:

$$\begin{aligned} K_0(J') + \epsilon K_1(J', t) &= H(\theta, J, t) + \frac{\partial S}{\partial t} \\ &= (H_0(J), + \epsilon H_1(\theta, J, t)) + \frac{\partial(\theta J' + \epsilon S_1)}{\partial t} \\ &= H_0\left(J' + \epsilon \frac{\partial S_1}{\partial \theta}\right) + \epsilon H_1(\theta, J', t) + \epsilon \frac{\partial S_1}{\partial t} \end{aligned}$$

Applying a 1st order Taylor expansion to H_0 :

$$\begin{aligned} H_0\left(J' + \epsilon \frac{\partial S_1}{\partial \theta}\right) &\approx H_0(J') + \left(J' + \epsilon \frac{\partial S_1}{\partial \theta} - J'\right) \frac{\partial H_0}{\partial J'} \\ &= H_0(J') + \epsilon \frac{\partial S_1}{\partial \theta} \frac{\partial H_0}{\partial J'} \end{aligned}$$

At this point we need to take a detour and prove the following simplification that $\omega_0(J') = \frac{\partial H_0(J')}{\partial J'}$. The canonical equation for the new angle variable θ' is:

$$\dot{\theta}' = \frac{\partial K}{\partial J'}$$

We define the unperturbed frequency ω_0 as the rate of change of the new angle in the limit where the perturbation vanishes. Using our identity $K_0 = H_0$ (from standard action-angle variables), we obtain:

$$\omega_0(J') \equiv \lim_{\epsilon \rightarrow 0} \dot{\theta}' = \frac{\partial K_0(J')}{\partial J'} = \frac{\partial H_0(J')}{\partial J'} \quad (31)$$

This proves that ω_0 is the frequency of the unperturbed system evaluated at the value of the new action J' . Therefore the 1st order Taylor approximation to the H_0 term is

$$H_0\left(J' + \epsilon \frac{\partial S_1}{\partial \theta}\right) \approx H_0(J') + \epsilon \omega_0(J') \frac{\partial S_1}{\partial \theta}$$

Continuing on to equating the expressions for $K(J', t)$:

$$K_0(J') + \epsilon K_1(J', t) = H_0(J') + \epsilon \left(\omega_0(J') \frac{\partial S_1}{\partial \theta} + H_1(\theta, J', t) + \frac{\partial S_1}{\partial t} \right)$$

Equating terms in powers of ϵ , for $O(1)$: $K_0 = H_0$, for $O(\epsilon)$:

$$K_1(J', t) = \omega_0 \frac{\partial S_1}{\partial \theta} + \frac{\partial S_1}{\partial t} + H_1(\theta, J', t) \quad (32)$$

We must solve for K_1 and S_1 . To ensure J' is a constant of motion K_1 must be independent of θ , we can decompose H_1 into a sum of its oscillating and fluctuating components: $H_1(\theta, J', t) = \langle H_1 \rangle_{J', t} + \tilde{H}(\theta, J', t)$, which requires a mini-proof. Starting from the transformed hamiltonian in Eq. (32), we apply the averaging operator $\langle \cdot \rangle = \frac{1}{2\pi} \int_0^{2\pi} (\dots) d\theta$ to both sides:

$$\langle K_1 \rangle = \omega_0 \left\langle \frac{\partial S_1}{\partial \theta} \right\rangle + \left\langle \frac{\partial S_1}{\partial t} \right\rangle + \langle H_1 \rangle$$

1. Left Side: Since K_1 is independent of θ by requirement, $\langle K_1 \rangle = K_1(J', t)$.

2. Periodicity Constraint: For the canonical transformation to be physically well-defined, the generating function S_1 must be 2π -periodic in θ . By the Fundamental Theorem of Calculus:

$$\left\langle \frac{\partial S_1}{\partial \theta} \right\rangle = \frac{1}{2\pi} [S_1(2\pi, J', t) - S_1(0, J', t)] = 0$$

3. Gauge Choice: We choose the purely time-dependent part of S_1 to be zero (or constant), such that $\langle S_1 \rangle = 0$. Consequently, $\langle \partial_t S_1 \rangle = \partial_t \langle S_1 \rangle = 0$.

Conclusion: The equation reduces to the first-order energy shift:

$$K_1(J', t) = \langle H_1(\theta, J', t) \rangle$$

Subtracting this average from the original equation defines the oscillating part \tilde{H}_1 , which drives the wiggles in the generating function:

$$\omega_0 \frac{\partial S_1}{\partial \theta} + \frac{\partial S_1}{\partial t} = -(H_1 - \langle H_1 \rangle) \equiv -\tilde{H}_1$$

therefore the 1st correction to the Hamiltonian is given by:

$$K(J', t) = H_0(J') + \langle H_1(\theta, J', t) \rangle \quad (33)$$

Using the Hamiltonian in Eq. (33) and the canonical position equation $\dot{\theta}' = \frac{\partial K(J', t)}{\partial J'}$ we can obtain the 1st order approximation to the frequency:

$$\begin{aligned} \dot{\theta}' &= \frac{\partial}{\partial J'} (H_0(J') + \langle H_1(\theta, J', t) \rangle) \\ &= \frac{\partial H_0(J')}{\partial J'} + \frac{\partial \langle H_1(\theta, J', t) \rangle}{\partial J'} \end{aligned}$$

Substituting $\omega_0(J') = \frac{\partial H_0(J')}{\partial J'}$ as in Eq.(31) we obtain:

$$\dot{\theta}' = \omega_0(J') + \frac{\partial \langle H_1(\theta, J', t) \rangle}{\partial J'}$$

Recall that because canonical transformations preserve the form of Hamiltons equations, this means that the canonical frequency relation is preserved: $w'(J', t) = \dot{\theta}'$. Thus the 1st order frequency is given by:

$$w'(J', t) = \omega_0(J') + \frac{\partial \langle H_1(\theta, J', t) \rangle}{\partial J'} \quad (34)$$

6.2.1 Perturbed Oscillator

The Hamiltonian for the simple harmonic oscillator with a quartic perturbation is given by

$$H(q, p) = \frac{p^2}{2m} + \frac{1}{2}m\omega_0^2 q^2 + \epsilon q^4$$

where the unperturbed part is given by $H_0 = \frac{p^2}{2m} + \frac{1}{2}m\omega_0^2 q^2$ and the perturbation is $H_1 = \epsilon q^4$. First we transformation to Action-Angle variables. We previously found in section 5 that the old coordinates are given by inverting the canonical transformations given by Eqs. (23, 22): $p(t) = \sqrt{2mE} \cos(\omega t + \phi)$, and $q(t) = \sqrt{\frac{2E}{m\omega^2}} \sin(\omega t + \phi)$. In section 6 we found the action angle variables for this system to be given by $J = E/\omega_0$ and $\theta = \omega t + \phi$. Combining these two results we obtain the transformations in terms of the action-angle coordinates:

$$p(t) = \sqrt{2m\omega_0 J} \cos(\theta), \quad q(t) = \sqrt{\frac{2J}{m\omega_0}} \sin(\theta) \quad (35)$$

Using the transformations ins Eq. (35), the unperturbed Hamiltonian in action-angle variables becomes $H_0(J) = \omega_0 J$, and the unperturbed Hamiltonian becomes

$$H_1(\theta, J) = \epsilon \left(\sqrt{\frac{2J}{m\omega_0}} \sin(\theta) \right)^4 = \epsilon \left(\frac{4J^4}{m^2\omega_0^2} \right) \sin^4(\theta)$$

The 1st order correction to the transformed Hamiltonian due to the perturbation is given by $K(J', t) = H_0(J') + \langle H_1(\theta, J', t) \rangle$ as in Eq. (33). The time-independent version is nearly identical just without any time dependence: $K(J) = H_0(J) + \langle H_1(\theta, J) \rangle$. Calculating $\langle H_1(\theta, J) \rangle$:

$$\begin{aligned}\langle H_1(\theta, J) \rangle &= \frac{1}{2\pi} \int_0^{2\pi} \epsilon \left(\frac{4J^4}{m^2\omega_0^2} \right) \sin^4(\theta) d\theta \\ &= \frac{3\epsilon J^2}{2m^2\omega_0^2}\end{aligned}$$

Therefore the 1st order correction to the Hamiltonian is given by $K(J) = H_0(J) + \langle H_1(\theta, J) \rangle = \omega_0 J + \frac{3\epsilon J^2}{2m^2\omega_0^2}$. Now that we have the Hamiltonian, we can easily obtain the perturbed frequency from Eq.(34):

$$\begin{aligned}\omega'(J) &= \frac{\partial K(J)}{\partial J} = \frac{\partial}{\partial J} \left(\omega_0 J + \frac{3\epsilon J^2}{2m^2\omega_0^2} \right) \\ &= \omega_0 + \frac{3\epsilon J}{m^2\omega_0^2}\end{aligned}\tag{36}$$

Thus, the 1st order correction to the frequency of oscillation due to a quartic perturbation is given by $\omega'(J) = \omega_0 + \frac{3\epsilon J}{m^2\omega_0^2}$.

6.2.2 Simple Pendulum

The Hamiltonian for the simple pendulum is given by

$$H = \frac{p^2}{2ml^2} + mgl(1 - \cos \theta)$$

At first glance there is no perturbation to work with, and we cannot treat $mgl(1 - \cos \theta)$ as the perturbation in this problem as it corresponds to a different scenario (spinning rotor). The idea is to approximate the behavior of the simple pendulum by applying the Taylor series expansion to the cosine term (which is what gives us all the trouble). Recall the series expansion of cosine: $\cos(\theta) = 1 - \theta^2/2 + \theta^4/4! - O(\theta^6)$, if we substitute this into the potential energy term, we get $mgl(1 - \cos \theta) \approx mgl\theta^2/2 - mgl\theta^4/24 - O(\theta^6)$. Up to 4th order, The Hamiltonian then becomes

$$H = \underbrace{\frac{p^2}{2ml^2}}_{H_0} + \underbrace{\frac{mgl\theta^2}{2} - \frac{mgl\theta^4}{24}}_{H_1}$$

This form looks exactly like the anharmonic oscillator with a quartic perturbation! Let $I = ml^2$, $\omega_0 = \sqrt{g/l}$, the Hamiltonian can be written as $H_0 = \frac{p^2}{2I} + \frac{I\omega_0^2\theta^2}{2}$ and $H_1 = -\frac{I\omega_0^2\theta^4}{24}$. This problem is slightly different with constants. Skipping the action-angle variables and the TIHJE, the transformations are:

$$p = \sqrt{2I\omega_0 J_0} \cos \phi, \quad \theta = \sqrt{\frac{2J_0}{m\omega_0}} \sin \phi$$

Under this canonical transformation, H_0 becomes $H_0 = \omega_0 J_0$, and H_1 becomes:

$$H_1(J_0, \omega_0) = -\frac{I\omega_0^2}{24} \left(\sqrt{\frac{2J_0}{m\omega_0}} \sin \phi \right)^4 = \frac{J_0^2}{6I} \sin^4 \phi$$

The average of the perturbed Hamiltonian over one cycle is given by:

$$\begin{aligned}\langle H_1(J_0, \omega_0) \rangle &= \frac{1}{2\pi} \int_0^{2\pi} \frac{J_0^2}{6I} \sin^4 \phi d\phi = -\frac{J_0^2}{12\pi I} (3/8) \\ &= -\frac{J_0^2}{16I}\end{aligned}$$

So the 1st order approximation to the transformed Hamiltonian is $K(J) = H_0(J) + \langle H_1(J, \omega_0) \rangle_{\omega_0} = \omega_0 J - \frac{J^2}{16I}$. Now we can use the canonical equation for the position coordinate to find the corrected frequency:

$$\begin{aligned} w'(J) &= \frac{\partial K}{\partial J} = \frac{\partial}{\partial J} \left(\omega_0 J - \frac{J^2}{16I} \right) \\ &= \omega_0 - \frac{J}{8I} \end{aligned}$$

we have therefore derived the first order (technically forth order in θ) correction to the frequency of the simple harmonic oscillator given by $w'(J) = \omega_0 - \frac{J}{8I}$.

6.2.3 Perturbed Kepler Orbits

The Hamiltonian for the perturbed 3D Kepler system is given by

$$H(q, p) = \frac{1}{2m} \left(p_r^2 + \frac{p_\theta^2}{r^2} + \frac{p_\phi^2}{r^2 \sin^2 \theta} \right) - \frac{k}{r} + \frac{\epsilon}{r^2} \quad (37)$$

where the unperturbed Hamiltonian is $H_0 = \frac{p_r^2}{2m} + \frac{L^2}{2mr^2} - \frac{k}{r}$ and the perturbation is $H_1 = \frac{\epsilon}{r^2}$, with ϵ being a small parameter. We first transform H_0 to Action-Angle variables. The three action variables are defined as:

$$\begin{aligned} J_\phi &= \frac{1}{2\pi} \oint p_\phi d\phi = L_z \\ J_\theta &= \frac{1}{2\pi} \oint p_\theta d\theta = \frac{1}{2\pi} \oint \sqrt{L^2 - \frac{J_\phi^2}{\sin^2 \theta}} d\theta = L - J_\phi \\ J_r &= \frac{1}{2\pi} \oint \sqrt{2mE + \frac{2mk}{r} - \frac{(J_\theta + J_\phi)^2}{r^2}} dr \end{aligned}$$

Evaluating the radial integral using the residue at infinity and the pole at $r = 0$, we obtain $J_r = -L + \frac{mk}{\sqrt{-2mE}}$. Summing these actions, we define the principal action $J = J_r + J_\theta + J_\phi = \frac{mk}{\sqrt{-2mE}}$. Solving for energy, the unperturbed Hamiltonian in action-angle variables becomes:

$$H_0(J) = -\frac{mk^2}{2J^2} \quad (38)$$

To find the 1st order correction to the Hamiltonian $K(J) = H_0(J) + \langle H_1 \rangle$, we average the perturbation over the unperturbed orbit. Using the property that for a Keplerian orbit $\langle 1/r^2 \rangle = \frac{m\omega_r}{L}$, where $\omega_r = \frac{mk^2}{J^3}$ and $L = J_\theta + J_\phi$:

$$\begin{aligned} \langle H_1(J_r, J_\theta, J_\phi) \rangle &= \frac{1}{(2\pi)^3} \int \frac{\epsilon}{r^2} d\theta_r d\theta_\theta d\theta_\phi = \epsilon \left\langle \frac{1}{r^2} \right\rangle \\ &= \frac{\epsilon m^2 k^2}{J^3(J_\theta + J_\phi)} \end{aligned} \quad (39)$$

The corrected Hamiltonian to first order is therefore:

$$K(J_r, J_\theta, J_\phi) = -\frac{mk^2}{2J^2} + \frac{\epsilon m^2 k^2}{J^3(J_\theta + J_\phi)} \quad (40)$$

We now obtain the perturbed frequencies $\omega'_i = \partial K / \partial J_i$. The radial frequency is:

$$\omega'_r = \frac{\partial K}{\partial J_r} = \frac{mk^2}{J^3} - \frac{3\epsilon m^2 k^2}{J^4(J_\theta + J_\phi)} \quad (41)$$

And the angular frequency (in the orbital plane) is:

$$\omega'_\theta = \frac{\partial K}{\partial J_\theta} = \frac{mk^2}{J^3} - \frac{3\epsilon m^2 k^2}{J^4(J_\theta + J_\phi)} - \frac{\epsilon m^2 k^2}{J^3(J_\theta + J_\phi)^2} \quad (42)$$

Because $\omega'_r \neq \omega'_\theta$, the degeneracy of the Kepler problem is broken. The difference between these frequencies gives the rate of perihelion precession:

$$\dot{\psi}_{prec} = \omega'_\theta - \omega'_r = -\frac{\epsilon m^2 k^2}{J^3(J_\theta + J_\phi)^2} \quad (43)$$

This result shows that the Action-Angle technique directly yields the physical precession rate of the orbit without requiring the integration of the equations of motion.

6.3 Adiabatic Invariance

The Adiabatic Theorem states that for a Hamiltonian $H(q, p, \lambda(t))$, the action variable J remains constant provided $\dot{\lambda}/\lambda \ll \omega$. To prove this, we consider a Hamiltonian where $\lambda(t)$ is a time dependent parameter that is low changing: $\dot{\lambda}/\lambda \ll \omega$. We perform a canonical transformation to (θ, J) variables using a type 2 generating function $W(q, J, \lambda)$:

$$K(\theta, J, t) = H(q, p, \lambda) + \frac{\partial W}{\partial t}$$

where the $\frac{\partial W}{\partial t}$ term becomes $\frac{\partial W(q, J, t)}{\partial t} = \frac{\partial W}{\partial q} \frac{\partial q}{\partial t} + \frac{\partial W}{\partial J} \frac{\partial J}{\partial t} + \frac{\partial W}{\partial \lambda} \frac{\partial \lambda}{\partial t}$, because (q, J, t) are independent of each other, their derivative wrt each other are zero, thus simplifying the expression to $\frac{\partial W(q, J, t)}{\partial t} = \dot{\lambda} \frac{\partial W}{\partial \lambda}$. Separately, because of the adiabatic condition of the slowly varying λ , the Hamiltonian is approximately the total energy $H(q, p, \lambda) = E(J, \lambda)$:

$$K(\theta, J, t) = E(J, \lambda) + \dot{\lambda} \frac{\partial W}{\partial \lambda} \quad (44)$$

Applying Hamilton's canonical equations for the canonical momentum::

$$\dot{J} = -\frac{\partial K}{\partial \theta} = -\dot{\lambda} \frac{\partial^2 W}{\partial \theta \partial \lambda} \quad (45)$$

In the adiabatic process, we are interested in the evolution J over a long period of time. To see if J invariant, we look at its change over one cycle of motion T :

$$\Delta J = \oint \dot{J} dt = \int_0^{2\pi} -\dot{\lambda} \frac{\partial^2 W}{\partial \theta \partial \lambda} dt$$

Since λ is slowly changing, we assume its $\dot{\lambda}$ is approximately constant over a cycle. We can also change variables using $dt = d\theta/\omega \approx d\theta/\omega$. The change in canonical momentum

$$\Delta J = \oint \dot{J} dt \approx -\frac{\dot{\lambda}}{\omega} \int_0^{2\pi} \frac{\partial}{\partial \theta} \left(\frac{\partial W}{\partial \lambda} \right) d\theta = -\frac{\dot{\lambda}}{\omega} \left[\frac{\partial W}{\partial \lambda} \right]_0^{2\pi} = 0$$

the generating function must preserve periodicity of the transformed Hamiltonian so it must be equal at the boundaries. Therefore $\Delta J = 0 \rightarrow J = \text{const.}$ up to 1st order $\dot{\lambda}$. The action variable is preserved, hence J is an adiabatic invariant.

6.3.1 Adiabatic Pendulum

A standard application of the adiabatic theorem is to consider the physical scenario of slowly changing the length of a simple pendulum, what happens to the energy and frequency of the system when the pendulum length is shortened or lengthened? The Hamiltonian of the system up to θ^2 approximation is given by:

$$H = \frac{p^2}{2mL^2} + \frac{mgL\theta^2}{2} = E_i$$

With $p = \sqrt{2mL^2(E_i - \frac{1}{2}mgL\theta^2)}$. If we slowly vary the length of the pendulum, such that $\dot{L}/L \ll \omega$, then the canonical momentum is constant $\Delta J = 0 \rightarrow J_i = J_f$. Thus our first step is to calculate the canonical momenta/action variable:

$$2\pi J = \oint p d\theta = \oint \sqrt{2mL^2(E - \frac{1}{2}mgL\theta^2)} d\theta$$

Applying a change of variables $\theta = q' \sin \phi \rightarrow d\theta = q' \cos \phi d\phi$, where $q' = \sqrt{\frac{2E}{mgL}}$. The limits of integration correspond to the turning points of oscillation at $p = 0$, so $\theta = q' \sin \phi = q' \rightarrow \phi = \pi/2$. The integral becomes:

$$\begin{aligned} 2\pi J &= 2 \int_{-\pi/2}^{\pi/2} \sqrt{2mL^2(E - \frac{1}{2}mgL(2E \sin^2 \phi / mgL))} q' \cos \phi d\phi \\ &= 2q' \sqrt{2mL^2 E} \int_{-\pi/2}^{\pi/2} \sqrt{1 - \sin^2 \phi} \cos \phi d\phi \rightarrow 2q' \sqrt{2mL^2 E} \underbrace{\int_{-\pi/2}^{\pi/2} \cos^2 \phi d\phi}_{= \pi/2} \\ &\rightarrow J = E \sqrt{\frac{L}{g}} = \frac{E}{\omega} \end{aligned}$$

Applying the adiabatic theorem, if the change in length is small enough then the canonical momentum is constant. Recall this implies $\Delta J = 0 \rightarrow J_i = J_f$. Using the expression derived from the momentum, we further obtain:

$$\begin{aligned} J_i &= J_f \rightarrow E_i \sqrt{\frac{L_i}{g}} = E_f \sqrt{\frac{L_f}{g}} \\ &\rightarrow E_f = E_i \sqrt{\frac{L_i}{L_f}} \end{aligned}$$

If $L_f = 2L_i$, then $E_f = E_i \sqrt{\frac{L_i}{2L_i}} \rightarrow E_f = E_i / \sqrt{2}$. So lengthening the length reduces the energy by $\sqrt{2}$.

If $L_f = L_i/2$, then $E_f = E_i \sqrt{\frac{L_i}{L_i/2}} \rightarrow E_f = E_i \sqrt{2}$. So shortening the length increases the energy by $\sqrt{2}$. For the changes in the frequencies, we know that:

$$\begin{aligned} J_i &= J_f \rightarrow \frac{E_i}{\omega_i} = \frac{E_f}{\omega_f} \\ \omega_f &= \omega_i \frac{E_f}{E_i} \end{aligned}$$

Using our previous results for the energies:

If $L_f = 2L_i$, then $E_f = E_i / \sqrt{2}$, which means the frequency changes by $\omega_f = \omega_i \frac{E_f}{E_i} \rightarrow \omega_i \frac{E_i / \sqrt{2}}{E_i} \rightarrow \omega_f = \omega_i / \sqrt{2}$. So lengthening the pendulum decreases the frequency.

If $L_f = L_i/2$, then $E_f = E_i \sqrt{2}$, which means the frequency changes by $\omega_f = \omega_i \frac{E_f}{E_i} \rightarrow \omega_i \frac{E_i \sqrt{2}}{E_i} \rightarrow \omega_f = \omega_i \sqrt{2}$. So shortening the pendulum increases the frequency.

6.3.2 Adiabatic Well

Imagine a 1D box with infinitely hard walls with a particle with energy $T = \frac{p^2}{2m}$, bouncing back forth at some frequency ω . Because the walls are perfectly hard, there is no loss of energy and momentum is conserved, so its momentum changes discontinuously when it bounces off the wall. The Hamiltonian for this system is given by

$$H = \frac{p^2}{2m} = E$$

Now imagine the the width of the well changes adiabatically from its original length L . How does energy and frequency of oscillation change if the L is changed in this manner? Because L changes adiabatically, it satisfies $\dot{L}/L \ll \omega$, which means the action is approximately constant through one

period: $\Delta J = 0 \rightarrow J_i = J_f$. The action variable is calculated as:

$$\begin{aligned} 2\pi J &= \oint p dx = \int_0^L p dx + \int_L^0 -p dx \rightarrow \int_0^L p dx + \int_0^L p dx \\ &= 2 \int_0^L p dx = 2 \int_0^L \sqrt{2mE} dx \\ J &= \frac{L\sqrt{2mE}}{\pi} \end{aligned}$$

Inverting $J = \frac{L\sqrt{2mE}}{\pi}$ in terms of E we obtain $E = \frac{\pi^2 J^2}{2mL^2}$, which looks exactly the energy levels in the infinite potential well for a quantum particle, except there is no quantization of course. To see how the energy changes, we utilize $J_i = J_f$ to derive the following:

$$\begin{aligned} J_i &= J_f \\ \frac{L_i \sqrt{2mE_i}}{\pi} &= \frac{L_f \sqrt{2mE_f}}{\pi} \\ E_f &= E_i \left(\frac{L_i}{L_f} \right)^2 \end{aligned}$$

If $L_f = 2L_i$, then $E_f = E_i/4$. So lengthening the length reduces the energy by 1/4.

If $L_f = L_i/2$, then $E_f = 4E_i$. So shortening the length increases the energy by 4. For the frequencies,

We can find the angular frequency ω from the canonical equation $\omega = \frac{\partial H}{\partial J} = \frac{dE(J,L)}{dJ}$:

$$\begin{aligned} \omega &= \frac{dE(J,L)}{dJ} = \frac{d}{dJ} \left(\frac{\pi^2 J^2}{2mL^2} \right) = \frac{\pi^2 J}{mL^2} \\ &= \frac{\pi}{L} \sqrt{\frac{2E}{m}} \end{aligned}$$

Inverting to find E in terms of ω : $E = \frac{m(\omega L)^2}{2L^2}$. Adapting the energy shift $E_f = E_i \left(\frac{L_i}{L_f} \right)^2$ in terms of ω we get:

$$\omega_f = \omega_i \left(\frac{L_i}{L_f} \right)^2$$

If $L_f = 2L_i$, then $\omega_f = \omega_i/4$. So lengthening the length reduces the frequency by 1/4.

If $L_f = L_i/2$, then $\omega_f = 4\omega_i$. So shortening the length increases the frequency by 4.

A Calculus Review

A.1 The Chain Rule for Change of Variables

If we have a function $f(q_1, \dots, q_n)$ and we perform a coordinate transformation to a new set Q_1, \dots, Q_n where $q_i = q_i(Q_1, \dots, Q_n)$, the partial derivative with respect to the new coordinates is given by:

$$\frac{\partial f}{\partial Q_j} = \sum_{i=1}^n \frac{\partial f}{\partial q_i} \frac{\partial q_i}{\partial Q_j}$$

In the context of Canonical Transformations $(q, p) \rightarrow (Q, P)$, this manifests in the preservation of the Poisson brackets.

A.2 Differentials of Multivariable Functions

For a function $f(x_1, \dots, x_n)$, the total differential is defined as:

$$df = \frac{\partial f}{\partial x_1} dx_1 + \frac{\partial f}{\partial x_2} dx_2 + \dots + \frac{\partial f}{\partial x_n} dx_n$$

In Hamilton-Jacobi theory, we often seek a generating function $S(q, P, t)$. Its total differential is:

$$dS = \sum_i \left(\frac{\partial S}{\partial q_i} dq_i + \frac{\partial S}{\partial P_i} dP_i \right) + \frac{\partial S}{\partial t} dt$$

By comparing this to the differential relations derived from the Legendre transformation (e.g., $p_i = \frac{\partial S}{\partial q_i}$), we derive the equations of motion for the system.

In mechanics, we often deal with functions defined along a path, $f(q_n, \dot{q}_n, t)$. The **total derivative** with respect to time accounts for both explicit time dependence and the implicit dependence through the coordinates:

$$\frac{df}{dt} = \frac{\partial f}{\partial t} + \sum_{i=1}^n \left(\frac{\partial f}{\partial q_i} \dot{q}_i + \frac{\partial f}{\partial \dot{q}_i} \ddot{q}_i \right)$$

A common source of confusion is the distinction between $\frac{\partial}{\partial t}$ (holding coordinates fixed) and $\frac{d}{dt}$ (following the system's trajectory).

A.3 Differentiation Under the Integral Sign (Leibniz Rule)

In the Principle of Least Action, we vary integrals of the form $S = \int_{t_1}^{t_2} L(q, \dot{q}, t) dt$. The general Leibniz rule for a parameter α is:

$$\frac{d}{d\alpha} \int_{a(\alpha)}^{b(\alpha)} f(x, \alpha) dx = f(b(\alpha), \alpha) \frac{db}{d\alpha} - f(a(\alpha), \alpha) \frac{da}{d\alpha} + \int_{a(\alpha)}^{b(\alpha)} \frac{\partial f}{\partial \alpha} dx$$

In fixed-endpoint variations, the first two terms vanish, allowing the δ operator to commute with the integral: $\delta \int L dt = \int \delta L dt$.

B Calculus of Variations

In the calculus of variations, the symbol δ represents a virtual displacement in the functional space. While the differential d represents a change along a specific path (a change in the independent variable t), the variation δ represents a change from one path to a neighboring path at a fixed value of t .

Definition of the Variation

Let $q(t)$ be the actual path and $q(t, \alpha)$ be a family of neighboring paths such that $q(t, 0) = q(t)$. The variation of the function q is defined as the differential with respect to the parameter α :

$$\delta q = \left(\frac{\partial q(t, \alpha)}{\partial \alpha} \right)_{\alpha=0} d\alpha$$

For a function $f(q, \dot{q}, t)$, the variation is defined as the first-order change induced by δq :

$$\delta f = f(q + \delta q, \dot{q} + \delta \dot{q}, t) - f(q, \dot{q}, t) \approx \frac{\partial f}{\partial q} \delta q + \frac{\partial f}{\partial \dot{q}} \delta \dot{q}$$

Algebraic Properties

The operator δ behaves like a derivative operator and satisfies the standard rules of linear calculus:

1. **Linearity:** $\delta(af + bg) = a\delta f + b\delta g$ for constants a, b .
2. **Product Rule:** $\delta(fg) = f\delta g + g\delta f$.
3. **Quotient Rule:** $\delta(f/g) = \frac{g\delta f - f\delta g}{g^2}$.
4. **Power Rule:** $\delta(f^n) = nf^{n-1}\delta f$.

Commutation Relations

A fundamental property used in the derivation of the Euler-Lagrange equations is that the variational operator δ commutes with both the differential d and the integral \int .

1. Commutation with Differentiation: The variation of a derivative is the derivative of the variation:

$$\delta \left(\frac{dq}{dt} \right) = \frac{d}{dt} (\delta q) \quad (46)$$

Proof sketch: Since t and α are independent variables, the mixed partial derivatives commute: $\frac{\partial^2 q}{\partial \alpha \partial t} = \frac{\partial^2 q}{\partial t \partial \alpha}$.

2. Commutation with Integration: For an integral with fixed limits $[t_1, t_2]$:

$$\delta \int_{t_1}^{t_2} L dt = \int_{t_1}^{t_2} \delta L dt \quad (47)$$

This allows us to move the δ operator inside the action integral, which is the starting point for Hamilton's Principle.

Comparison: δ vs. d

It is important to distinguish between the total differential df and the variation δf :

- $df = \frac{\partial f}{\partial t} dt + \sum \frac{\partial f}{\partial q_i} dq_i$: Represents the change in f as the system moves an infinitesimal distance dt **along the path**.
- $\delta f = \sum \frac{\partial f}{\partial q_i} \delta q_i$: Represents the change in f at a **fixed time** t due to a change in the **choice of path**.

Note that $\delta t = 0$ in standard fixed-time variations, whereas $dt \neq 0$ for the evolution of the system.

C Alternative Derivations:

C.1 Hamilton's Equations

We shall derive Hamilton's equation with a variational approach. The Lagrangian in terms of the Hamiltonian is: $L = \sum_i p_i \dot{q}_i - H$ And Hamilton's principle states:

$$\delta S = \int_a^b \delta L dt = 0$$

Substituting $L = \sum_i p_i \dot{q}_i - H$ into Hamilton's Principle:

$$\begin{aligned}
\delta S &= \int_a^b \delta \left(\sum_i p_i \dot{q}_i - H(q, p, t) \right) dt = 0 \\
&= \sum_i \int_a^b dt \left(\dot{q}_i \delta p_i + p_i \delta \dot{q}_i - \delta q_i \frac{\partial H}{\partial q_i} - p_i \frac{\partial H}{\partial p_i} \right) = 0 \\
&= \sum_i \int_a^b dt \left(\left(\dot{q}_i - \frac{\partial H}{\partial p_i} \right) \delta p_i + \left(-\delta q_i \frac{\partial H}{\partial q_i} \right) \right) + \sum_i \int_a^b dt (\delta q_i \dot{p}_i) = 0 \\
&= \sum_i \int_a^b dt \left(\left(\dot{q}_i - \frac{\partial H}{\partial p_i} \right) \delta p_i + \left(-\delta q_i \frac{\partial H}{\partial q_i} \right) \right) - \sum_i \int_a^b dt (\dot{p}_i \delta q_i) = 0 \\
&= \sum_i \int_a^b dt \left(\left(\dot{q}_i - \frac{\partial H}{\partial p_i} \right) \delta p_i + \left(-\dot{p}_i - \frac{\partial H}{\partial q_i} \right) \delta q_i \right) = 0
\end{aligned}$$

For any arbitrary variation of each $\delta q_i(t)$ and $\delta p_i(t)$, $\delta S = 0$ only if each of the following equations are satisfied for all i :

$$\dot{q}_i = \frac{\partial H}{\partial p_i}, -\dot{p}_i = \frac{\partial H}{\partial q_i} \quad (48)$$

D Useful Resources

[Dr. Jorge Diaz' YouTube Channel](#): Incredibly detailed theoretical physics channel at the graduate level, offers very insightful videos into profound concepts in physics, this was one of my main sources.

[Michael Fowler's graduate level lectures](#): I found these lecture notes to be incredibly insightful and helpful for understanding canonical transformations and Hamilton Jacobi theory. He follows Landau, but offers more detail and explanation, this was one of my main sources.

[Carlos Gutierrez' Classical Mechanics](#): Incredibly detailed and thorough, I had an easier time understanding the logic for the canonical transformations and the Hamilton Jacobi theory, this was one of my main sources. I downloaded all his lecture notes but the notes are currently taken down.

[MITOCW Classical Mechanics](#): Lecture notes on canonical transformations, Hamilton Jacobi theory and angle action variables were good supplementary and references, but not nearly as detailed as Dr. Fowler's notes, but still was one of my main sources.

[Noethers Theorem by Rubens](#): One of the most helpful sources I found for understanding Noether's theorem.