Symmetry and conservation laws in Classical Mechanics

What I understood

Abhishek Daniel D

Lets make a step by step progression from elementary considerations to comprehensive formalism on what we know about conservation laws in Newtonian mechanics:

(i) Newton's first law of motion states that, *An object will remain in the state of rest or of uniform motion in a straight line unless a non uniform external force acts on it.* Note that here the concept of inertial frames of reference, uniform flow of absolute time and the validity of Euclidean geometry of physical 3D space are all assumed.
In the absence of external force and considering the space is homogeneous, the momentum of an object is the constant of motion(COM).

(ii) Now moving to the system of two particles, we have the Newton's second law of motion as our equation(EOM). Together with his third law, we can related the forces the particle exert on each other:

$$\vec{p_1} = m_1\dot{\vec{r_1}}\ ,\ \vec{p_2} = m_2\dot{\vec{r_2}}$$
$$\vec{F_{12}} = \dot{\vec{p_1}}\ ,\ \vec{F_{21}} = \dot{\vec{p_2}}$$

$$\vec{F_{12}} = -\vec{F_{21}} \implies \vec{F_{12}} + \vec{F_{21}} = 0 \implies \dot{\vec{p_1}} + \dot{\vec{p_2}} = 0$$

Thus after doing integration we realize that $(\vec{p_1} + \vec{p_2})$ is conserved i.e. it is a COM.

(iii) If these two forces arise from an inter-particle potential, we have

$$\vec{F_{12}} = -\nabla_1 V(\vec{r_1} - \vec{r_2}) \ , \ \vec{F_{21}} = -\nabla_2 V(\vec{r_1} - \vec{r_2})$$

The third law holds good because of the translation invariance of the potential. Thus we can see that conservation of linear momentum is connected to a symmetry called translation symmetry.

In the previous slide we introduced translation symmetry. Let's discuss what exactly is it. Consider a vector potential $V(x) = x\hat{x}$. Let's take a line $x = a$, every point in that line has same potential. If we move every point a little by $x = a + \delta x$, every point will experience equal amount of potential (equipotential surface, well it is in 3D case, a line or curve in 2D representation of this surface). Thus there exist certain quantities here as invariant. This is called *translation invariance.*

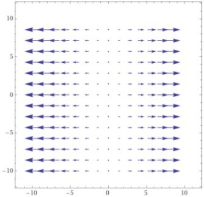


Figure: $V(x) = x\hat{x}$

Consider a system with n degrees of freedom described by generalized coordinates $q_r = q(q_1, q_2, q_3, ..., q_n)$. We can define the Lagrangian $L(q, \dot{q})$ for the system as,

$$L = T - V \tag{1}$$

where T is the kinetic energy and V is the potential energy. Note that we can have explicit time dependence in L.

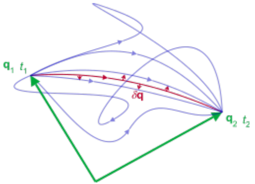
We are assuming that the **Hessian matrix**(H) whose $(ij)^{th}$ elements are defined as $\frac{\partial^2 L}{\partial \dot{q_i} \partial \dot{q_j}}$ to be non-singular i.e. $det(H) \neq 0$. So that we can convert Lagrangian to Hamiltonian and vice versa.

Principle of least action

Action: It is a numerical description of what path it takes to reach its final state $q(t_2)$ from the initial state. It is denoted as S. It is a functional quantity i.e. $S \to S[q(t)]$. It is defined as

$$S[q(t)] = \int_{t_1}^{t_2} L(q, \dot{q})dt \tag{2}$$

According to least action principle, the actual path it takes from $q(t_1)$ and $q(t_2)$ is the path of action which is minimum. i.e. $\delta S = 0$



We already derived the Euler Lagrangian equation from D'Alembret's principle. Let us derive the same from the principle of least action.

$$\begin{aligned}
\delta S &= S[q(t)+\delta q(t)] - S[q(t)] \\
&= \int_{t_1}^{t_2} L(q+\delta q, \dot{q}+\delta\dot{q})dt - \int_{t_1}^{t_2} L(q,\dot{q})dt \\
&= \int_{t_1}^{t_2} \left[L(q+\delta q, \dot{q}+\delta\dot{q}) - L(q,\dot{q})\right] dt \\
&= \int_{t_1}^{t_2} \left[(T(\dot{q}+\delta\dot{q}) - V(q+\delta q)) - (T(\dot{q}) - V(q))\right] dt
\end{aligned}$$

Euler Lagrangian equation(contd.)

Using Taylor approximation, $f(x+\delta x) \approx f(x) + \frac{\partial f(x)}{\partial x}\delta x$

$$
\begin{aligned}
\delta S &= \int_{t_1}^{t_2} \left[T(\dot{q}) + \frac{\partial T}{\partial \dot{q}}\delta\dot{q} - V(q) - \frac{\partial V}{\partial q}\delta q - T(\dot{q}) + V(q) \right] dt \\
&= \int_{t_1}^{t_2} \left[\frac{\partial T}{\partial \dot{q}}\delta\dot{q} - \frac{\partial V}{\partial q}\delta q \right] dt
\end{aligned}
$$

Consider $\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}}\delta q\right) = \frac{\partial T}{\partial \dot{q}}\delta\dot{q} + \frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}}\right)\delta q$

Thus, $\frac{\partial T}{\partial \dot{q}}\delta\dot{q} = \frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}}\delta q\right) - \frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}}\right)\delta q$

$$
\delta S = \int_{t_1}^{t_2} \left[\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}}\delta q\right) - \frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}}\right)\delta q - \frac{\partial V}{\partial q}\delta q \right] dt
$$

Since path starts from $\delta q(t_1)$ to $\delta q(t_2)$ the variations at those points are zero i.e. $\delta q(t_1) = \delta q(t_2) = 0$

Thus,

$$\delta S = -\int_{t_1}^{t_2} \left[\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}}\right) + \frac{\partial V}{\partial q} \right] \delta q dt$$

From principle of least action, $\delta S = 0$. Therefore,

$$\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}}\right) + \frac{\partial V}{\partial q} = 0$$

Since T is independent of q and V is independent of $\dot{q}$, we can write the above equation in terms of Lagrangian as

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}}\right) - \frac{\partial L}{\partial q} = 0 \tag{3}$$

The equation(3) holds good for n number of independent variables. Therefore generally,

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_r}\right) - \frac{\partial L}{\partial q_r} = 0 \qquad (4)$$

where $r = 1, 2, 3, ..., n$. Equation(4) is called the *Euler Lagrangian equation.*

In Lagrangian framework, a COM is any function $f(q, \dot{q}, t)$ whose total time derivative vanishes

$$\begin{aligned} EOM &\implies \frac{d}{dt} f(q, \dot{q}, t) = 0 \\ &\implies f(q, \dot{q}, t) = COM \end{aligned}$$

Right now there is no link to symmetry. But this fact is to be remembered.

A particular generalized coordinate is called a cyclic coordinate if the that coordinate doesn't explicitly appear in the Lagrangian. For example, let's say that the Lagrangian of a system as no dependence on a particular coordinate, say, q_k but depend on $\dot{q_k}$, then the coordinate q_k is considered as 'cyclic'.

Therefore the corresponding Euler Lagrangian equation at the coordinate q_k is,

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q_k}}\right) = 0$$

Thus,

$$p_k = \frac{\partial L}{\partial \dot{q_k}} \tag{5}$$

Here equation (5) is a COM.

We can express this as an invariant quantity of the Lagrangian. The cyclic nature of q_k means that the Lagrangian is invariant under infinitesimal translation or the shift in q_k which is independent of time.

Thus, If $\delta q_k = \epsilon \delta_{rk} \implies \delta L = 0$, then $p_k =$ COM.

Where ϵ is the infinitesimal time-independent parameter so that $\delta \dot{q_k} = 0$ and *kronecker delta* will make sure that only q_k th coordinate alone shifted.

Also in case of Lagrangian $L(q, \dot{q}, t)$, if the time doesn't appear explicitly, then we say that **'time is cyclic'**. So now we can say that our Lagrangian is invariant under *time translation*.

Thus if the time is cyclic in Lagrangian $L(q_i, \dot{q}_i, t)$ i.e. $\frac{\partial L}{\partial t} = 0$ then the total time derivative of L is

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

$$\frac{d}{dt}\left(\sum_i \frac{\partial L}{\partial \dot{q}_r}\dot{q}_r - L(q,\dot{q},t)\right) = 0$$

$$\implies \sum_i \frac{\partial L}{\partial \dot{q}_r}\dot{q}_r - L(q,\dot{q},t) = \mathrm{COM} \implies \sum_i p_r\dot{q}_r - L = \mathrm{COM}$$

This is nothing but the Hamiltonian obtained from the **Legendre Transformation** of the Lagrangian.

Note

This is not actually Hamiltonian of a system which is defined by $H = T + V$. This is called **energy function**. It is numerically equal to Hamiltonian when there is no energy loss or in other words no work is extracted from the system.

To define Legendre Transformation, we must understand what is convex function. A function $f(x)$ whose interval I is defined to be smooth(no holes i.e. continuous function & atleast 2 derivatives exist) and real valued i.e. $f : I \in \mathbb{R}$ is said to convex if and only if the second derivative of the function is strictly greater than zero i.e. $f''(x) > 0$.

This definition of convex function holds good only when the function has one variable. If a given function has multiple independent variable, the concept of convexity is kinda absurd and abstract.

In higher dimensional case a smooth and real valued function $f(x_1, x_2, ..., x_n)$ is said to be convex iff it's Hessian matrix is non-singular i.e. $det(H) \neq 0$ where $H_{ij} = \frac{\partial^2 f}{\partial x_i \partial x_j}$. Since Lagrangian is assumed to be satisfy this condition it is convex function.

For a given convex function $f(x)$, we can define a new variable s such that $s = f'(x)$ so that $f(x) \mapsto F(s)$ as $F(s) = sx - f(x)$.

Since $f(x)$ is monotonic function, it is invertible i.e. $x = (f')^{-1}(s)$. Thus,

$$F(s) = sx(s) - f(x(s)) \tag{6}$$

For example we can use Legendre transformation to map/change from one thermodynamical variable to other like internal energy $U(V, S) \mapsto$ Helmholtz free energy$F(V, T)$ and Helmholtz free energy $F(V, T) \mapsto$ Gibbs free energy $G(P, T)$.

An infinitesimal point transformation in Lagrangian mechanics is a small change made in each generalized coordinate at each time as $\delta q_r = \varepsilon \phi(q_r)$ where r goes from 1 to n. Here ε is the infinitesimal time dependent parameter. Also, we can say $\delta \dot{q_r} = \varepsilon \frac{d\phi(q_r)}{dt} = \varepsilon \frac{\partial \phi(q_r)}{\partial q_l} \dot{q}_l$

Consider that the Lagrangian is invariant under this transformation. Therefore, $\delta L = 0$

$$\begin{aligned} \delta L &= \frac{\partial L}{\partial q_r} \delta q_r + \frac{\partial L}{\partial \dot{q_r}} \delta \dot{q_r} = 0 \\ \implies & \frac{dp_r}{dt} \delta q_r + p_r \frac{d\delta q_r}{dt} = 0 \implies \frac{d}{dt}(p_r \delta q_r) = 0 \\ & G(q,p) = p_r \phi(q_r) = \text{COM} \end{aligned} \tag{7}$$

Here $G(q,p)$ is called the Generic COM as fn. in phase space. This is so called the Generator of the transformation.

An Hamiltonian H of a system is defined as $H = T + V$ as we said previous this nothing the total energy of the system in phase space. Of course for a energy conserved system Hamiltonian is related to Lagrangian by Legendre transformation as $H(q, p, t) = \sum(p_i \dot{q}_i) - L(q, \dot{q}, t)$

Now consider the total differential of the Hamiltonian,

$$dH(q, p, t) = d(p\dot{q}) - dL(q, \dot{q}, t)$$

$$\frac{\partial H}{\partial q}dq + \frac{\partial H}{\partial p}dp + \frac{\partial H}{\partial t}dt = pd\dot{q} + \dot{q}dp - \frac{\partial L}{\partial q}dq - \frac{\partial L}{\partial \dot{q}}d\dot{q} - \frac{\partial L}{\partial t}dt$$

$$\frac{\partial H}{\partial q}dq + \frac{\partial H}{\partial p}dp + \frac{\partial H}{\partial t}dt = pd\dot{q} + \dot{q}dp - \dot{p}dq - pd\dot{q} - \frac{\partial L}{\partial t}dt$$

From the last equation, we can compare the coefficients of differentials on both sides. Thus,

$$dq : \frac{\partial H}{\partial q} = -\dot{p} \tag{8}$$

$$dp : \frac{\partial H}{\partial p} = \dot{q} \tag{9}$$

$$dt : \frac{\partial H}{\partial t} = -\frac{\partial L}{\partial t} \tag{10}$$

Here equations (8) and (9) are collectively known as Hamilton's equations.

It worth noting that for a system of n degree of freedom the Euler Lagrange equation gives n 2nd order PDE while using Hamilton's equations we get $2n$ 1st order PDE.

Previously we defined $p_i = \frac{\partial L}{\partial \dot{q}_i}$ it is obvious that p is nothing but the momentum of the system by doing the dimensional analysis. But this is more than that. This is called Conjugate momentum. This quantity is dependent on Lagrangian.

For a free particle in Cartesian coordinate Lagrangian is $L = \frac{1}{2}m(\dot{x}^2 + \dot{y}^2)$, here the conjugate momentum $p_x = m\dot{x}$ and $p_y = m\dot{y}$. This is linear momentum.

For a free particle in polar coordinate Lagrangian is $L = \frac{1}{2}m(\dot{r}^2 + r^2\dot{\theta}^2)$, here the conjugate momentum $p_\theta = mr^2\dot{\theta} = mr\omega$. This is angular momentum.

For a charged particle moving in a EM field, the conjugate momentum $\mathbf{p} = m\dot{\mathbf{v}} + q\mathbf{A}$, here the $\mathbf{p}$ have not only mechanical momentum also momentum from the electromagnetic vecor potential $\mathbf{A}$.

For given two functions $f(q,p)$ and $g(q,p)$ the Poisson bracket is defined as,

$$\{f,g\} = \frac{\partial f}{\partial q}\frac{\partial g}{\partial p} - \frac{\partial g}{\partial q}\frac{\partial f}{\partial p} \tag{11}$$

The Poisson brackets satisfies $\{q_r, q_s\} = 0$, $\{p_r, p_s\} = 0$ and $\{q_r, p_s\} = \delta_{rs}$ Note that the Poisson bracket is anti-commutative thus $\{q_r, p_s\} = -\{p_r, q_s\}$. Also it satisfies bilinearity and Jacobi identity

For a function $f(q,p,t)$ to be a COM $\frac{d}{dt}f(q,p,t) = 0$. Those functions of the dynamical variables which remain constant during the motion of the system are called integrals of the motion.

$$\frac{df}{dt} = \frac{\partial f}{\partial q}\dot{q} + \frac{\partial f}{\partial p}\dot{p} + \frac{\partial f}{\partial t} = 0$$

$$\frac{\partial f}{\partial q}\frac{\partial H}{\partial p} - \frac{\partial f}{\partial p}\frac{\partial H}{\partial q} + \frac{\partial f}{\partial t} = 0 \implies \{f,H\} + \frac{\partial f}{\partial t} = 0$$

In the case of Lagrangian Mechanics, let's say that we are going to transform $(q, \dot{q}, t) \mapsto (s, \dot{s}, t)$ Thus the Lagrangian changes $L(q, \dot{q}, t) \mapsto L_s(s, \dot{s}, t)$ and related by just replacement i.e. $L(q, \dot{q}, t) = L_s(s(q,t), \dot{s}(q, \dot{q}, t), t)$ then we can see that the form of Euler Lagrangian equation is invariant.

This is because in the configuration space when we transform $(q, \dot{q}) \mapsto (s, \dot{s})$, $\dot{s}$ is not allowed to transform arbitrarily.

But in the case of Hamiltonian mechanics, the Hamilton's equations are not invariant in such transformation. This is because in phase space, p can transform to some other P easily in mapping $(q, p) \mapsto (Q, P)$. Thus the form of Hamilton's equations are not preserved in general.

But under certain transformation, the form of the Hamilton's equations are invariant.These set of transformations are called **Canonical Transformation(CT)**.

Say we transform $(q, p) \mapsto (Q, P)$ and Q and P are defined such that $Q_i = Q_i(q, p)$ and $P_i = P_i(q, p)$. This is a canonical transformation if there exist $K(Q, P, t)$ such that,

$$\frac{\partial K}{\partial P_i} = \dot{Q}_i$$

$$\text{And} \quad \frac{\partial K}{\partial Q_i} = -\dot{P}_i$$

i.e. it leaves the form of the Hamilton's equations unchanged.

Note that here the Hamiltonian H is mapped to some function K i.e. $H(q, p, t)(Q, P, t)$. It is not necessary that H should be equal to K. The point of finding a different coordinate is it will given lesser number of PDEs to solve.

In Infinitesimal Canonical Transformation(ICT) Q and P differ infinitesimally from q and p:

$$Q_r = q_r + \delta q_r \; ; \; P_r = p_r + \delta p_r$$

In infinitesimal point transformation we set $\delta q_r = \varepsilon\phi(q_r)$ and $G(q,p) = p_r\phi(q_r)$ in Poisson bracket format we can write δq_r as:

$$\delta q_r = \varepsilon\{q_r, G(q,p)\} \tag{12}$$

Similarly, we can write δp_r as:

$$\delta p_r = \delta\left(\frac{\partial L}{\partial \dot{q}_r}\right) = \varepsilon\frac{\partial^2 L}{\partial \dot{q}_r \partial q_s}\phi_s(q) + \varepsilon\frac{\partial^2 L}{\partial \dot{q}_r \partial \dot{q}_s}\frac{\partial \phi_s(q)}{\partial q_l}\dot{q}_l$$

To further simply we must consider the assumption made in infinitesimal point transformation $\delta L = 0$

$$\delta L = \varepsilon \frac{\partial L}{\partial q_s}\phi(q_s) + \varepsilon \frac{\partial L}{\partial \dot{q_s}} \frac{\partial \phi(q_s)}{\partial q_l}\dot{q}_l = 0$$

Now differentiate the previous equation with respect to $\dot{q_r}$.
Thus we will get,

$$\varepsilon \frac{\partial^2 L}{\partial \dot{q_r} \partial q_s}\phi_s(q) + \varepsilon \frac{\partial^2 L}{\partial \dot{q_r} \partial \dot{q_s}} \frac{\partial \phi_s(q)}{\partial q_l}\dot{q}_l + \varepsilon p_s \frac{\partial \phi_s(q)}{\partial q_l} = 0$$

$$\delta p_r + \varepsilon p_s \frac{\partial \phi_s(q)}{\partial q_l} = 0$$

$$\implies \delta p_r = -\varepsilon p_s \frac{\partial \phi_s(q)}{\partial q_l} = \varepsilon \{p_r, G(q,p)\} \tag{13}$$

In equations (12) and (13), Note that we end up having a generator/generic COM which we obtained by imposing infinitesimal point transformation on the Lagrangian $G(q,p) = p_r\phi(q_r)$ which in turn proved as COM by imposing Euler Lagrange equation (our EOM) in the infinitesimal change in the Lagrangian.

Thus we can say,

(1) We needed EOM (Euler Lagrange equation) only to show that generator of CT is a COM.

(2) The infinitesimal point transformation symmetry of the Lagrangian appears as an ICT in phase space associated with a generator as its COM.

Initial we considered that in the infinitesimal point transformation the infinitesimal change(δ-variation) in q's as $\delta q_r = \phi_r(q)$ now with that knowledge lets extend that to even more generalized version by permitting the velocity dependence in ϕ i.e. $\delta q_r = \phi_r(q, \dot{q})$

Thus we have

$$\delta \dot{q}_r = \varepsilon \frac{d}{dt}\phi_r(q, \dot{q}) = \varepsilon \left(\frac{\partial \phi_r}{\partial q_r}\dot{q}_r + \frac{\partial \phi_r}{\partial \dot{q}_r}\ddot{q}_r \right)$$

Here we can see that the $\delta \dot{q}_r$ have an acceleration term. This makes the infinitesimal change in Lagrangian not equal to zero i.e. $\delta L \neq 0$.

Rather we consider, Under these changes the Lagrangian changes by a total time derivative of the some function i.e.

$$\delta L = \varepsilon \frac{d}{dt} F(q, \dot{q})$$

where F(q,) is a function of the given variables that is local in time. If it so then we have a dynamical symmetry of the Lagrangian.

Thus we get,

(1) $G(q, p) = p_r \phi_r(q, \dot{q}) - F(q, \dot{q})$ as the COM

(2) $\delta q_r = \varepsilon\{q_r, G(q, p)\}$ and $\delta p_r = \varepsilon\{p_r, G(q, p)\}$ still holds good in this larger dynamical symmetry.

Thus, The most general Infinitesimal symmetry of a Lagrangian is a dynamical symmetry where the generator of a CT being the COM.

Thank you!!!