# Introduction to the Calculus of Variations
# Lecture Notes
# Lecture 4

Swarnendu Sil

Spring 2021, IISc

# Classical Methods

## Symmetry and Noether's theorem

We have already seen that $H$ is a conserved quantity if $H$ does not depend explicitly on $t$. This is not a coincidence, but is just an example of a profound general fact.

$$\text{Symmetries of } \mathcal{L} \text{ or } H \quad \Leftrightarrow \quad \text{first integral.}$$

*Any symmetry of the Lagrangian (and thus also of the Hamiltonian and vice versa) corresponds to a first integral, i.e. a conserved quantity.*

This deep result is the Noether's theorem, named after its discoverer, the brilliant Emmy Noether. Conservation laws are ubiquitous in nature and so are symmetries. This theorem connects the two and tells us that these two apparently disconnected aspects of a physical system, symmetries and conserved quantities, are in a precise sense one and the same thing.

**Time translation symmetry** Now let us show in which sense no explicit $t$-dependence of the the Hamiltonian (and thus the Lagrangian too) is a symmetry. This symmetry is called the *time translation symmetry* or in other words, *invariance under time translations*. Consider the *one-parameter family of diffeomorphisms*

$$\phi_\tau(t) = t + \tau \qquad \text{for } \tau \in \mathbb{R}.$$

Thus, the curve $u : [a, b] \to \mathbb{R}^N$ is transformed to $u \circ \phi_\tau^{-1} : [a + \tau, b + \tau] \to \mathbb{R}^N$. Now, since the Hamiltonian does not depend explicitly on $t$, the same must be true for the Lagrangian density $f$. So we can write

$$f = f\left(u, \xi\right).$$

Now, it is clear that

$$\phi_\tau^{-1}\left(s\right) = s - \tau \qquad \text{for every } \tau \in \mathbb{R}.$$

Thus, we have

$$\frac{d}{ds}\left[u \circ \phi_\tau^{-1}\right](s) = \frac{d}{ds}\left[u\left(s - \tau\right)\right] = \dot{u}\left(s - \tau\right).$$

Hence, for any $\theta \in (a, b]$, we have

$$\begin{aligned}
&\int_{a+\tau}^{\theta+\tau} f\left(u \circ \phi_\tau^{-1}(s), \frac{d}{ds}\left[u \circ \phi_\tau^{-1}\right](s)\right) \mathrm{d}s \\
&\qquad = \int_{a+\tau}^{\theta+\tau} f\left(u(s-\tau), \dot{u}(s-\tau)\right) \mathrm{d}s \\
&\qquad = \int_{a}^{\theta} f\left(u(t), \dot{u}(t)\right) \mathrm{d}t. \qquad (1)
\end{aligned}$$

Thus, we have the invariance

$$I_\theta[u] = I_\theta[u \circ \phi_\tau^{-1}] \qquad \text{for any } \theta \in (a, b].$$

To understand better what is happening, perhaps it is instructive to compare (1) to the following string of equalities

$$\begin{aligned}
I_\theta[u \circ \phi_\tau^{-1}] &= \int_{a+\tau}^{\theta+\tau} f\left(s, u \circ \phi_\tau^{-1}(s), \frac{d}{ds}\left[u \circ \phi_\tau^{-1}\right](s)\right) \mathrm{d}s \\
&= \int_{a+\tau}^{\theta+\tau} f\left(s, u(s-\tau), \dot{u}(s-\tau)\right) \mathrm{d}s \\
&= \int_{a}^{\theta} f\left(t+\tau, u(t), \dot{u}(t)\right) \mathrm{d}t \neq I_\theta[u],
\end{aligned}$$

which would have resulted instead if $f$ were to depend explicitly on the $t$ variable.

**Invariance** Now we are going to define more precisely what we mean by a symmetry. We begin with the definition of a $C^r$-smoothly varying one parameter family of diffeomorphisms.

**Definition 1** ($C^r$-smoothly varying one parameter family of diffeomorphisms)**.** *Let $X, Y$ be smooth manifolds and let $r, p \geq 1$ be integers. A* $C^r$-smoothly varying one parameter family of $C^p$-diffeomorphisms on $X$ *is a map*

$$\Phi : \mathbb{R} \times X \to Y$$

*satisfying the following properties.*

- ***one parameter family of $C^p$-diffeomorphisms** For each $s \in \mathbb{R}$, the map*

  $$\Phi(s, \cdot) := \phi_s(\cdot) : X \to \phi_s(X) \subset Y$$

  *is a diffeomorphism of $X$ (onto its image) into $Y$ of class $C^p$.*

- ***$C^r$-smoothly varying** The map*

  $$s \mapsto \Phi(s, \cdot) := \phi_s(\cdot),$$

*considered as a map from $\mathbb{R}$ into $C^p(X;Y)$ is of class $C^r$. If we only write diffeomorphisms instead of $C^p$-diffeomorphisms, it is understood that we are talking about $C^\infty$-diffeomorphisms. Likewise, if we only write smoothly varying instead of $C^r$-smoothly varying, this means the map*

$$s \mapsto \Phi(s, \cdot) := \phi_s(\cdot),$$

*considered as a map from $\mathbb{R}$ into $C^p(X;Y)$ is $C^\infty$.*

Now we can define the notion of invariance under the action of a family of diffeomorphisms.

**Definition 2** (Invariance)**.** *Let $\phi_s : [a,b] \times \mathbb{R}^N \to \mathbb{R} \times \mathbb{R}^N$, $s \in \mathbb{R}$ be a $C^2$-*smoothly varying one-parameter family of diffeomorphisms*, which can be written as*

$$\phi_s(t,z) = \left(\phi_s^0(t), \bar{\phi}_s(z)\right) \quad \text{for every } t \in [a,b] \text{ and for every } z \in \mathbb{R}^N,$$

*such that*

$$\phi_0(t,z) = (t,z) \qquad \text{for every } t \in [a,b] \text{ and for every } z \in \mathbb{R}^N.$$

*A Lagrangian is invariant under the action of the family of diffeomorphisms $\{\phi_s\}_{s\in\mathbb{R}}$ if it satisfies*

$$\int_{\phi_s^0(a)}^{\phi_s^0(\theta)} f\left(t_s, \left[\bar{\phi}_s \circ u \circ \left(\phi_s^0\right)^{-1}\right](t_s), \frac{d}{dt_s}\left[\bar{\phi}_s \circ u \circ \left(\phi_s^0\right)^{-1}\right](t_s)\right) \,\mathrm{d}t_s$$
$$= \int_a^\theta f(t, u(t), \dot{u}(t)) \,\mathrm{d}t \qquad \text{for every } \theta \in (a,b],$$

*for every $s \in \mathbb{R}$, where $t_s = \phi_s^0(t)$.*

## Noether's theorem

**Theorem 3** (Noether's theorem)**.** *Let $f \in C^2\left([a,b] \times \mathbb{R}^N \times \mathbb{R}^N\right)$. Suppose the Lagrangian*

$$I[u] = \int_a^b f(t, u(t), \dot{u}(t)) \,\mathrm{d}t$$

*is invariant under the action of the family of diffeomorphisms $\{\phi_s\}$ as above. Then the following expression*

$$\left\langle f_\xi(t, u(t), \dot{u}(t)), \left.\frac{d}{ds}\left[\bar{\phi}_s(u(t))\right]\right|_{s=0} \right\rangle$$
$$+ \left[f(t, u(t), \dot{u}(t)) - \langle f_\xi(t, u(t), \dot{u}(t)), \dot{u}(t)\rangle\right] \left.\frac{d}{ds}\left[\phi_s^0(t)\right]\right|_{s=0}$$

*is* ***constant*** *along any solution $u(t)$ of the EL equations for $I$, i.e. defines a* ***first integral****.*

*Proof.* First we prove a special case where the $t$ variable is unchanged, i.e. $\phi_s^0(t) = t$ for every $t \in [a,b]$ for every $s \in \mathbb{R}$. Thus, the invariance condition reduces to

$$\int_a^\theta f\left(t, \left[\bar{\phi}_s \circ u\right](t), \frac{d}{dt}\left[\bar{\phi}_s \circ u\right](t)\right) \,\mathrm{d}t$$
$$= \int_a^\theta f(t, u(t), \dot{u}(t)) \,\mathrm{d}t$$

for every $s \in \mathbb{R}$, for every $\theta \in (a,b]$ and for any $u \in C^2\left([a,b];\mathbb{R}^N\right)$. [1]

Fix $\theta \in (a,b]$. Now differentiating with respect to $s$ and using the fact that $\bar{\phi}_0$ is identity on $\mathbb{R}^N$, we deduce, by invariance,

$$\begin{aligned}
0 &= \frac{d}{ds}\left(\int_a^\theta f\left(t, \left[\bar{\phi}_s \circ u\right](t), \frac{d}{dt}\left[\bar{\phi}_s \circ u\right](t)\right) \,\mathrm{d}t\right)\Bigg|_{s=0} \\
&= \int_a^\theta \left\langle f_u(t, u(t), \dot{u}(t));\, \frac{d}{ds}\left[\bar{\phi}_s \circ u\right]\Big|_{s=0}(t)\right\rangle \mathrm{d}t \\
&\quad + \int_a^\theta \left\langle f_\xi(t, u(t), \dot{u}(t));\, \frac{d}{dt}\left[\frac{d}{ds}\left[\bar{\phi}_s \circ u\right]\Big|_{s=0}(t)\right]\right\rangle \mathrm{d}t \\
&= \int_a^\theta \left\langle \frac{d}{dt}\left[f_\xi(t, u(t), \dot{u}(t))\right];\, \frac{d}{ds}\left[\bar{\phi}_s \circ u\right]\Big|_{s=0}(t)\right\rangle \mathrm{d}t \\
&\quad + \int_a^\theta \left\langle f_\xi(t, u(t), \dot{u}(t));\, \frac{d}{dt}\left[\frac{d}{ds}\left[\bar{\phi}_s \circ u\right]\Big|_{s=0}(t)\right]\right\rangle \mathrm{d}t
\end{aligned}$$

In the last line we substituted for $f_u$ using the EL equations. So, we obtained so far

$$\begin{aligned}
0 &= \int_a^\theta \left\langle \frac{d}{dt}\left[f_\xi(t, u(t), \dot{u}(t))\right];\, \frac{d}{ds}\left[\bar{\phi}_s \circ u\right]\Big|_{s=0}(t)\right\rangle \mathrm{d}t \\
&\quad + \int_a^\theta \left\langle f_\xi(t, u(t), \dot{u}(t));\, \frac{d}{dt}\left[\frac{d}{ds}\left[\bar{\phi}_s \circ u\right]\Big|_{s=0}(t)\right]\right\rangle \mathrm{d}t \\
&= \int_a^\theta \frac{d}{dt}\left[\left\langle f_\xi(t, u(t), \dot{u}(t));\, \frac{d}{ds}\left[\bar{\phi}_s \circ u\right]\Big|_{s=0}(t)\right]\right\rangle \mathrm{d}t
\end{aligned}$$

[1] Note that a consequence of invariance is that if $u$ satisfies the EL equations, so does $\bar{\phi}_s \circ u$ for every $s \in \mathbb{R}$. Thus, we have,

$$\frac{d}{dt}\left[f_\xi\left(t, \left[\bar{\phi}_s \circ u\right](t), \frac{d}{dt}\left[\bar{\phi}_s \circ u\right](t)\right)\right] = f_u\left(t, \left[\bar{\phi}_s \circ u\right](t), \frac{d}{dt}\left[\bar{\phi}_s \circ u\right](t)\right)$$

for every $s \in \mathbb{R}$ and for every $t \in (a,\theta)$. We would not use this.

Thus, we have

$$\left\langle f_\xi\left(\theta,u\left(\theta\right),\dot u\left(\theta\right)\right);\frac{d}{ds}\left[\bar\phi_s\circ u\right]\Big|_{s=0}\left(\theta\right)\right\rangle$$
$$=\left\langle f_\xi\left(a,u\left(a\right),\dot u\left(a\right)\right);\frac{d}{ds}\left[\bar\phi_s\circ u\right]\Big|_{s=0}\left(a\right)\right\rangle.$$

Since $\theta\in(a,b]$ is arbitrary, this proves the special case.

Now we are going to prove the general case. We want to reduce it to the special case we just proved. So we artifically introduce a variable $\tau$ and consider $t$ as a new dependent variable on the same footing as $u$ to transform the problem to the previous case, but on $\mathbb{R}^{N+1}$ instead.

For $\tau\in[a,b]$, we write $t=t\left(\tau\right):=\tau$ and set the new Lagrangian density

$$\begin{aligned}\bar f\left(t\left(\tau\right),\ u\left(t\left(\tau\right)\right),\ \frac{dt}{d\tau},\ \frac{d}{d\tau}\left[u\left(t\left(\tau\right)\right)\right]\right)\\ :=f\left(t,u\left(t\right),\frac{\frac{d}{d\tau}\left[u\left(t\left(\tau\right)\right)\right]}{\frac{dt}{d\tau}}\right)\frac{dt}{d\tau}\\ =f\left(t,u\left(t\right),\dot u\left(t\right)\right)\frac{dt}{d\tau}.\end{aligned}$$

So the Lagrangian

$$\bar I\left(t,u\right):=\int_{\tau_0}^{\tau_1}\bar f\left(t\left(\tau\right),\ u\left(t\left(\tau\right)\right),\ \frac{dt}{d\tau},\ \frac{d}{d\tau}\left[u\left(t\left(\tau\right)\right)\right]\right)\ \mathrm{d}\tau$$

does not depend explicitly on $\tau$ anymore.

Thus we can apply the previous result to $\bar I$ and deduce that

$$\begin{aligned}\left\langle\bar f_{\bar\xi}\left(t\left(\tau\right),\ u\left(t\left(\tau\right)\right),\ \frac{dt}{d\tau},\ \frac{d}{d\tau}\left[u\left(t\left(\tau\right)\right)\right]\right);\frac{d}{ds}\left[\bar\phi_s\circ u\right]\Big|_{s=0}\left(t\left(\tau\right)\right)\right\rangle\\ +\bar f_{\xi^0}\left(t\left(\tau\right),\ u\left(t\left(\tau\right)\right),\ \frac{dt}{d\tau},\ \frac{d}{d\tau}\left[u\left(t\left(\tau\right)\right)\right]\right)\frac{d}{ds}\left[\phi_s^0\circ t\right]\Big|_{s=0}\left(\tau\right)\end{aligned}$$

is a first integral, where $\xi^0$ is the $dt/d\tau$ variable and $\bar\xi$ stands for the $d\left[u\left(t\left(\tau\right)\right)\right]/d\tau$ variable. Now from the definition of $\bar f$, clearly at $s=0$, we have

$$\bar f_{\bar\xi}=f_\xi$$

and we compute

$$\bar f_{\xi^0}=f-\left\langle f_\xi;\dot u\right\rangle.\quad\left(\text{ since }\phi_s^0\left(t\left(\tau\right)\right)=t\left(\tau\right)=\tau,\text{ we have }dt/d\tau=1\text{ at }s=0\right)$$

This proves the theorem. □

**Examples of conservation laws**

Now we illustrate Noether's theorem by some simple examples.

**Example 1: Linear momentum conservation** Let $m > 0$ be the mass and $x(t) \in \mathbb{R}^3$ be the position of a point particle. Let the potential energy function $U : \mathbb{R}^3 \to \mathbb{R}$ be independent of the $x_3$ coordinate, i.e.

$$U(x) = U(x_1, x_2).$$

The Lagrangian density, as usual, is

$$f(t, x, \xi) = \frac{1}{2} m\xi^2 - U(x).$$

It is easy to check that the family of transformations

$$h_s(x) := x + se_3 = (x_1, x_2, x_3 + s), \qquad s \in \mathbb{R}$$

leaves the Lagrangian invariant. Now Noether's theorem tells us that the third component of **linear momemtum**

$$p_3 := m\dot{x}_3$$

is a first integral. Similarly, if $U \equiv 0$, then $m\dot{x}$ is a first integral.

**Example 2: Angular momentum conservation** Once again consider the same Lagrangian density. But this time let the potential energy function $U : \mathbb{R}^3 \to \mathbb{R}$ satisfy

$$U(x) = U(R_3[\phi] x) \quad \text{for all } \phi \in [0, 2\pi],$$

where the **rotation matrix** $R_3[\phi]$ represents a rotation about the $x_3$-axis through an angle $\phi$ and its precise expression is

$$R_3[\phi] = \begin{pmatrix} \cos\phi & -\sin\phi & 0 \\ \sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Clearly, if $\{\phi_s\}_{s\in\mathbb{R}}$ is a one parameter family of angles, and satisfies $\phi_0 = 0$, then the family of transformations

$$h_s(x) := R_3[\phi_s] x, \qquad s \in \mathbb{R}$$

leaves the Lagrangian invariant. Now one can check that Noether's theorem tells us that the third component of **angular momemtum**

$$M_3 := (u \wedge m\dot{x}) \cdot e_3$$

is a first integral, where $\wedge$ is the vector product ( cross product ).

**Conjugate momemta and cyclic variables** The first of our example can be generalized in a sense. Assume we are given a Hamiltonian which has no explicit dependence, i.e.

$$H\left(t,u,v\right)=H\left(u,v\right).$$

The variables $u$ and $v$ in the arguments of the Hamiltonian are in a sense **conjugate variables**. For each $1 \leq i \leq N$, the variable $v_i$, is called the **conjugate momenta** of the variable $u_i$ (for the same $i$).

Suppose also that for some $1 \leq i \leq N$, we have

$$\frac{\partial H}{\partial u_i}=0.$$

Then the variable $u_i$ is called a **cyclic variable**. Now it can be shown that the Noether's theorem implies that **the conjugate momemta for a cyclic variable is a first integral**. More precisely,

$$\frac{\partial H}{\partial u_i}=0 \quad \Rightarrow \quad v_i \text{ is a first integral.}$$

This actually can be extended to Hamiltonians with explicit dependence on $t$ as well. We have already seen that if $t$ is a 'cyclic variable', i.e. if $\frac{\partial H}{\partial t}=0$, then $H$ itself is a first integral.

Conjugate momenta of a variable can be and usually is defined in terms of the Lagrangian.

**Definition 4** (Conjugate momenta)**.** *Let $f \in C^1\left([a,b]\times\mathbb{R}^N\times\mathbb{R}^N\right)$, $f = f\left(t,u,\xi\right)$ be a given Lagrangian density. For $1 \leq i \leq N$, the function $f_{\xi_i}$ is called the **conjugate momentum** associated to the variable $u_i$.*

**Remark 5.** *The variables $u_i$, $1 \leq i \leq N$, are often called* position *or* generalized position *variables. This is the reason why the Lagrangian density is often written as*

$$f=f\left(t,q,\dot{q}\right)$$

*in the physics literature, where $q_i$s are the genenalized position or coordinate variables and*

$$p_i:=\frac{\partial f}{\partial \dot{q}_i}$$

*is the momentum or generalized momentum conjugate to $q_i$.*

As we have already seen in the discussion of Legendre transform and the Hamiltonian, the variable $v$ in the Hamiltonian basically represents $f_\xi$. So it is clear why we call $v_i$ the momentum conjugate to $u_i$ in terms of the Hamiltonian. However, note that only the $u_i$ variables have conjugate momenta. Neither the $\dot{u}_i$ variables in the Lagrangian formulation nor the $v_i$ variables in the Hamiltonian formulation has any conjugate momenta. Also note that there is a distinction between the *velocity* $\dot{q}$ and the *momentum* $f_\xi$, which might appear surprising due to the high-school physics maxim 'momentum = mass × velocity' and mass being a constant number. To appreciate fully the conceptual difference between the two, interested reader can try to right down the Lagrangian density for a relativistic particle moving in the Minkowski 4-space and calculate the momenta and the velocities. This exercise would also shed some light on why $H$ should be considered the 'momentum' conjugate to $t$ variable. The actual definition of a cyclic variable is also in terms of the Lagrangian.

**Definition 6** (Cyclic variables)**.** *Let $f \in C^1\left([a,b] \times \mathbb{R}^N \times \mathbb{R}^N\right)$, $f = f(t,u,\xi)$ be a given Lagrangian density. For $1 \leq i \leq N$, the variable $u_i$ is called a cyclic variable for the Lagrangian density $f$ if the Lagrangian density does not depend explicitly on $u_i$. More precisely, if we have*

$$\frac{\partial f}{\partial u_i} = 0.$$

Once again, it is not hard to see why $\frac{\partial H}{\partial u_i} = 0$ would imply that $u_i$ is a cyclic variable.