Calculus of Variations and Noether's Theorem

Jamie Merchant 5597613

July 11, 2026

Contents

1 Motivation

Many problems in mathematics and physics can be thought of as optimisation over entire paths or shapes according to some criterion rather than individual points. Some natural examples include:

- Finding the shortest path between two points (geodesics).
- Determining the shape of a hanging chain (the catenary) that minimises potential energy.
- Determining the surface of revolution of minimal area between two fixed points (the catenoid).

In this essay we aim to build a framework for solving this type of optimisation problem and explore how we may re-frame classical mechanics as a variational calculus problem, allowing us to explore how certain symmetries of systems lead to the conservation laws. This gives us a more mathematically rigorous basis for much of physics previously known only empirically.

2 Variational Principles

2.1 Functionals and Admissible Curves

Quantities that depend not just on numbers but on entire functions or paths are known as *functionals* and assign a single real number to each function (or curve) within a function space. In this sense, a functional can be thought of as a 'function of a function'.
Below is a motivating example from [3, pp.40-42] of a functional we will return to:

Example 2.1.1 (Minimal Surface Area of Revolution)**.** Here we aim to minimise the surface area of revolution for a shape between two fixed points a and b:

$$A[y] = \int_a^b 2\pi y \sqrt{1 + \left(\frac{dy}{dx}\right)^2}\, dx.$$

Definition 2.1.2 (Functional)**.** [2, p.1,p.4] A *functional* is a mapping from a vector space of functions to the real numbers:

$$I : \mathcal{F} \to \mathbb{R},$$

where $\mathcal{F}$ denotes the function space.

In this essay we focus on a particularly important class of functionals in physics[1] called *actions*, which assign a number to an entire trajectory of a

[1]Similar functionals arise in other contexts, such as the surface area functional described above.

system over time. For a curve

$$\mathbf{q} \in C^2([t_1, t_2]; \mathbb{R}^n) = \mathbf{C}^2[t_1, t_2]$$

an action functional is defined by $I : \mathbf{C^2}[t_1, t_2] \to \mathbb{R}$

$$I[\mathbf{q}] = \int_{t_1}^{t_2} L(t, \mathbf{q}, \dot{\mathbf{q}})\, dt \tag{1}$$

where $\mathbf{q}(t) = (q_1(t), q_2(t), \ldots, q_n(t))$ is a vector of generalized coordinates, and the *Lagrangian* $L : [t_1, t_2] \times \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ is assumed to have continuous second-order partial derivatives with respect to t, q_i, and $\dot{q}_i$ for $i = 1, \ldots, n$. These assumptions will be explained as we proceed.

Definition 2.1.3 (Space of allowed paths)**.** [6, p.62] The *space of allowed paths* S consists of all curves in $\mathbf{C^2}[t_1, t_2]$ satisfying fixed endpoint conditions:

$$S = \{\mathbf{q} \in \mathbf{C^2}[t_1, t_2] : \mathbf{q}(t_1) = \mathbf{q}_1 \text{ and } \mathbf{q}(t_2) = \mathbf{q}_2\}$$

where $\mathbf{q}_1, \mathbf{q}_2 \in \mathbb{R}^n$ are given fixed endpoint vectors.

Definition 2.1.4 (Space of allowed variations)**.** [6, p.61] The *space of allowed variations* H is the set of curves in $\mathbf{C^2}[t_1, t_2]$ that vanish at the endpoints:

$$H = \{\mathbf{h} \in \mathbf{C^2}[t_1, t_2] : \mathbf{h}(t_1) = \mathbf{h}(t_2) = \mathbf{0}\}$$

With these definitions, the problem of finding an extremal trajectory becomes the task of determining which curve $\mathbf{q} \in S$ makes the functional stationary under all admissible variations $\mathbf{h} \in H$.

2.2 Differentiation of Functionals

To find which path or curve extremises the action and later understand how it behaves under transformations, we need a notion of differentiation in function spaces.

Given a curve $\mathbf{q} \in S$ and a variation $\mathbf{h} \in H$, we consider a perturbation

$$\mathbf{q}(t) \to \mathbf{q}(t) + \epsilon\mathbf{h}(t),$$

where ϵ is a small parameter. Since $\mathbf{h}(t_1) = \mathbf{h}(t_2) = 0$, the perturbed curve satisfies the same endpoint conditions as $\mathbf{q}$.

The *Gâteaux Derivative* captures the leading order change of the functional under such perturbations[2].

[2]Note that could use the Fréchet derivative however for our purposes this is sufficient.

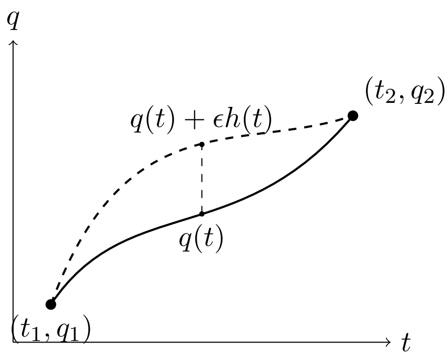


Figure 1: A variation of the path $q(t)$ with fixed endpoints.

Definition 2.2.1 (Gâteaux Derivative – First Variation)**.** [4][pp.328-329] Let X be a normed vector space and $F : X \to \mathbb{R}$ be a functional. We say F is Gâteaux differentiable at $x \in X$ if, for every direction of variation $h \in X$, the limit

$$D_h F[x] = \lim_{\epsilon \to 0} \frac{F(x + \epsilon h) - F(x)}{\epsilon} \text{ exists.}$$

Equivalently,

$$F(x + \epsilon h) - F(x) = \epsilon D_h F[x] + o(\epsilon) \quad \text{as } \epsilon \to 0,$$

where $\epsilon \in \mathbb{R}$ is the scalar parameter controlling the size of variation along h.

The map $h \mapsto D_h F[x]$ is called the *Gâteaux derivative* (or *first variation*[3]) of F at x.

The following proof of the differentiability of the action has been heavily adapted and generalised to $\mathbf{C^2}[t_1, t_2]$ from [1, pp.56-57]

Theorem 2.2.2 (Differentiability of the Action)**.** *The action functional* $I : \mathbf{C^2}[t_1, t_2] \to \mathbb{R}$ $I[\mathbf{q}] = \int_{t_1}^{t_2} L(t, \mathbf{q}, \dot{\mathbf{q}})\, dt$ *is Gâteaux differentiable with respect to its path* $\mathbf{q} \in S \subseteq \mathbf{C^2}[t_1, t_2]$ *and its derivative (first variation) along variation* $\mathbf{h} \in \mathbf{C^2}[t_1, t_2]$ *is given by*

$$D_{\mathbf{h}} I[\mathbf{q}] = \int_{t_1}^{t_2} \left(\sum_{i=1}^{n} \left(\frac{\partial L}{\partial q_i} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}_i} \right) h_i(t) \right) dt + \left[\sum_{i=1}^{n} \frac{\partial L}{\partial \dot{q}_i} h_i(t) \right]_{t_1}^{t_2} \tag{2}$$

Restricting now to $\mathbf{h} \in H$ *:*

$$D_{\mathbf{h}} I[\mathbf{q}] = \int_{t_1}^{t_2} \left(\sum_{i=1}^{n} \left(\frac{\partial L}{\partial q_i} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}_i} \right) h_i(t) \right) dt \tag{3}$$

[3] In other sources, $D_h F[x]$ is often written as $\delta F[x; h]$.

Proof. Let $\mathbf{h} \in \mathbf{C^2}[t_1, t_2]$ be an admissible variation and $\epsilon \in \mathbb{R}$ be small. Consider the perturbed path $\mathbf{q} + \epsilon\mathbf{h}$. Then

$$I[\mathbf{q} + \epsilon\mathbf{h}] - I[\mathbf{q}] = \int_{t_1}^{t_2} \Big(L(t, \mathbf{q} + \epsilon\mathbf{h}, \dot{\mathbf{q}} + \epsilon\dot{\mathbf{h}}) - L(t, \mathbf{q}, \dot{\mathbf{q}}) \Big)\, dt$$

Using the Taylor expansion of L to the first order in ϵ gives

$$L(t, \mathbf{q} + \epsilon\mathbf{h}, \dot{\mathbf{q}} + \epsilon\dot{\mathbf{h}}) - L(t, \mathbf{q}, \dot{\mathbf{q}}) = \epsilon \left(\nabla_{\mathbf{q}} L \cdot \mathbf{h} + \nabla_{\dot{\mathbf{q}}} L \cdot \dot{\mathbf{h}} \right) + o(\epsilon),$$

Substituting this into the integral

$$\begin{aligned} I[\mathbf{q} + \epsilon\mathbf{h}] - I[\mathbf{q}] &= \epsilon \int_{t_1}^{t_2} \left(\nabla_{\mathbf{q}} L \cdot \mathbf{h} + \nabla_{\dot{\mathbf{q}}} L \cdot \dot{\mathbf{h}} \right) dt + o(\epsilon) \\ &= \epsilon \int_{t_1}^{t_2} \sum_{i=1}^{n} \left(\frac{\partial L}{\partial q_i} h_i(t) + \frac{\partial L}{\partial \dot{q}_i} \dot{h}_i(t) \right) dt + o(\epsilon). \end{aligned}$$

Applying integration by parts component-wise[4]

$$\int_{t_1}^{t_2} \frac{\partial L}{\partial \dot{q}_i} \dot{h}_i(t)\, dt = \left[\frac{\partial L}{\partial \dot{q}_i} h_i(t) \right]_{t_1}^{t_2} - \int_{t_1}^{t_2} \frac{d}{dt} \frac{\partial L}{\partial \dot{q}_i} h_i(t)\, dt.$$

Combining terms

$$I[\mathbf{q} + \epsilon\mathbf{h}] - I[\mathbf{q}] = \epsilon \int_{t_1}^{t_2} \sum_{i=1}^{n} \left(\frac{\partial L}{\partial q_i} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}_i} \right) h_i(t)\, dt + \epsilon \sum_{i=1}^{n} \left[\frac{\partial L}{\partial \dot{q}_i} h_i(t) \right]_{t_1}^{t_2} + o(\epsilon) \tag{4}$$

Dividing by ϵ and taking $\epsilon \to 0$, we obtain the Gâteaux derivative along the variation $\mathbf{h}$:

$$D_{\mathbf{h}} I[\mathbf{q}] = \int_{t_1}^{t_2} \sum_{i=1}^{n} \left(\frac{\partial L}{\partial q_i} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}_i} \right) h_i(t)\, dt + \sum_{i=1}^{n} \left[\frac{\partial L}{\partial \dot{q}_i} h_i(t) \right]_{t_1}^{t_2}$$

Restricting now to $\mathbf{h} \in H$, since $\mathbf{h}(t_1) = \mathbf{h}(t_2) = 0$, we have

$$D_{\mathbf{h}} I[\mathbf{q}] = \int_{t_1}^{t_2} \sum_{i=1}^{n} \left(\frac{\partial L}{\partial q_i} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}_i} \right) h_i(t)\, dt.$$

□

[4] By the chain rule the term $\frac{d}{dt}\frac{\partial L}{\partial \dot{q}} = \frac{\partial^2 L}{\partial t \partial \dot{q}_i} + \sum_{j=1}^{n} \frac{\partial^2 L}{\partial q_j \partial \dot{q}_i} \dot{q}_j + \sum_{j=1}^{n} \frac{\partial^2 L}{\partial \dot{q}_j \partial \dot{q}_i} \ddot{q}_j$. This is why we need L to be second order partially differentiable and $\mathbf{q} \in \mathbf{C^2}[t_1, t_2]$.

2.3 Extremals and the Euler-Lagrange Equations

Before giving formal definitions, let us motivate the concept of an extremal. An extremal is a function that makes a functional stationary, analogous to a critical point of a real-valued function in finite dimensions. Just as a function $f : \mathbb{R}^n \to \mathbb{R}$ satisfies

$$\partial_{\mathbf{v}} f(\mathbf{a}) = 0 \quad \forall \mathbf{v} \in \mathbb{R}^n \implies \nabla f(\mathbf{a}) = 0,$$

an extremal of a functional F is a path where the first variation vanishes for all variations.

Definition 2.3.1 (Extremal of a functional)**.** [1, p.57] Let X be a vector space and let $I : X \to \mathbb{R}$ be a Gâteaux differentiable functional. A point $x \in X$ is called an *extremal* of F if

$$D_h F[x] = 0 \quad \text{for all } h \in X.$$

Before applying this idea to the expression for the first variation of the action

$$D_{\mathbf{h}} I[\mathbf{q}] = \int_{t_1}^{t_2} \sum_{i=1}^{n} \left(\frac{\partial L}{\partial q_i} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}_i} \right) h_i(t)\, dt,$$

we need a key result known as the *Fundamental Lemma of the Calculus of Variations*. This lemma allows us to conclude that when the integral of a function multiplied by a function that vanishes at the endpoints, then the function vanishes pointwise.

Lemma 2.3.2 (Fundamental Lemma of the Calculus of Variations)**.** *[2, p.9]. If $f(x)$ is continuous on $[x_1, x_2]$, and if*

$$\int_{x_1}^{x_2} f(x) h(x)\, dx = 0$$

for every function $h(x) \in C[x_1, x_2]$ such that $h(x_1) = h(x_2) = 0$, then $f(x) = 0$ for all $x \in [x_1, x_2]$.

Proof. [2, p.9] Suppose without loss of generality there exists a point $x_0 \in (x_1, x_2)$ such that $f(x_0) > 0$ (otherwise consider $-f$). Then since f is continuous there exists an interval $(a, b) \subset (x_1, x_2)$ such that $f(x) > 0$ for all $x \in (a, b)$.
If we set

$$h(x) = \begin{cases} 0 & x \in [x_1, a) \\ (x-a)(b-x) & x \in [a, b] \\ 0 & x \in (b, x_2] \end{cases}$$

then clearly $h(x)$ satisfies the lemma's conditions, however since $f(x)(x - a)(b-x) > 0$ on (a,b), we have

$$\int_{x_1}^{x_2} f(x)h(x)\,dx = \int_a^b f(x)h(x)\,dx > 0,$$

contradicting our assumption that $\int_{x_1}^{x_2} f(x)h(x)\,dx = 0$. Therefore, no such x_0 exists, and $f(x) = 0$ for all $x \in [x_1, x_2]$. $\square$

Once we have the Fundamental Lemma, we may closely follow the approach in [6, pp.61-62] to apply the lemma to each term in the sum of the first variation by considering component-wise variations

$$\mathbf{h}(t) = (0, \ldots, h_i(t), \ldots, 0) \in H,$$

for which $D_{\mathbf{h}}I[\mathbf{q}] = 0$. Applying the lemma to each component then yields the *Euler-Lagrange equations* for extremal curves.

Theorem 2.3.3 (Euler-Lagrange Equations)**.** *[6, p.62] Let $I : \mathbf{C^2}[t_1, t_2] \to \mathbb{R}$ be the action functional of the form*

$$I[\mathbf{q}] = \int_{t_1}^{t_2} L(t, \mathbf{q}, \dot{\mathbf{q}})\,dt$$

where $\mathbf{q}(t) = (q_1(t), q_2(t), \ldots, q_n(t))$, and L has continuous second-order partial derivatives with respect to t, x_i, and $\dot{x}_i$, $i = 1, 2, \ldots, n$. If $\mathbf{q}$ is an extremal for I in the space of allowed paths, S, then

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_i} - \frac{\partial L}{\partial q_i} = 0 \quad \textit{for all } i = 1, \ldots, n.$$

Proof. Suppose that the action functional I is extremal at a path $\mathbf{q} \in S$. By definition of an extremal, the first variation of I at $\mathbf{q}$ vanishes for every variation $\mathbf{h} \in \mathbf{C^2}[t_1, t_2]$, hence since $H \subset \mathbf{C^2}[\mathbf{t_1}, \mathbf{t_2}]$,

$$D_{\mathbf{h}}I[\mathbf{q}] = 0 \quad \text{for all } \mathbf{h} \in H.$$

From the differentiability of the action, the first variation is given by

$$D_{\mathbf{h}}I[\mathbf{q}] = \int_{t_1}^{t_2} \sum_{i=1}^{n} \left(\frac{\partial L}{\partial q_i} - \frac{d}{dt}\frac{\partial L}{\partial \dot{q}_i} \right) h_i(t)\,dt$$

where $\mathbf{h}(t_1) = \mathbf{h}(t_2) = \mathbf{0}$.

Since this integral equals zero for all $\mathbf{h} \in H$, we may consider variations purely in the ith component, $\mathbf{h} = (0, \ldots, h_i(t), \ldots, 0)$ for every $i = 1, \ldots, n$ giving us

$$\int_{t_1}^{t_2} \left(\frac{\partial L}{\partial q_i} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}_i} \right) h_i(t)\, dt = 0$$

where $h_i(t_1) = h_i(t_2) = 0$.

The Fundamental Lemma of the Calculus of Variations then implies that

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{q}_i} - \frac{\partial L}{\partial q_i} = 0 \quad \text{for all } i = 1, \ldots, n.$$

□

We now demonstrate the power of the Euler-Lagrange equations using a classical extremal problem:

Example 2.3.4 (Minimal Surface Area of Revolution - Continued)**.** [3, pp.40-42] To minimise[5] $A[y]$, take $L(y, y') = 2\pi y \sqrt{1 + (y')^2}$.

Applying the Euler-Lagrange equation:

$$\frac{d}{dx} \frac{\partial L}{\partial y'} - \frac{\partial L}{\partial y} = 0$$

and simplifying (algebra and integration omitted) yields the ODE

$$y' = \sqrt{\frac{y^2}{k^2} - 1}.$$

Solving by separation of variables gives

$$y = k \cosh\left(\frac{x - c}{k} \right).$$

This curve corresponds physically to the profile of a soap film spanning two circular rings, since surface tension minimises area.

2.4 Variable Endpoints and Free Boundary Conditions

So far, we restricted attention to variations $\mathbf{h} \in H = \{\mathbf{h} \in \mathbf{C^2}[t_1, t_2] : \mathbf{h}(t_1) = \mathbf{h}(t_2) = \mathbf{0}\}$ that vanish at endpoints, ensuring the boundary term in the first variation disappears, fixing the initial and final points of the trajectory. Inspired but heavily adapted and generalised from [6, pp.144-150] now remove this restriction, allowing the endpoints to vary.

Let the action functional be

$$I[\mathbf{q}] = \int_{t_1}^{t_2} L(t, \mathbf{q}, \dot{\mathbf{q}})\, dt$$

[5] It is worth noting that to confirm the curve we find is minimal we can take the second variation, analogous to the second derivative although this is omitted here.

where $\mathbf{q} \in \mathbf{C^2}[t_1 + \epsilon\tau_1, t_2 + \epsilon\tau_2]$. We consider a perturbation $\mathbf{h} \in \mathbf{C^2}[t_1 + \epsilon\tau_1, t_2 + \epsilon\tau_2]$

$$\mathbf{q} \to \mathbf{q}(t) + \epsilon\mathbf{h}(t),$$

and now allow the endpoints to vary

$$t_k \to t_k + \epsilon\tau_k, \qquad \mathbf{q}(t_k) \to \mathbf{q}(t_k) + \epsilon\boldsymbol{\eta}_k, \quad k = 1, 2.$$

The action evaluated on the perturbed path becomes

$$I[\mathbf{q} + \epsilon\mathbf{h}] = \int_{t_1+\epsilon\tau_1}^{t_2+\epsilon\tau_2} L(t, \mathbf{q}(t) + \epsilon\mathbf{h}, \dot{\mathbf{q}} + \epsilon\dot{\mathbf{h}})\, dt.$$

As before, we compute the first variation using the Gâteaux derivative

$$D_{\mathbf{h}} I[\mathbf{q}] = \lim_{\epsilon\to 0} \frac{I[\mathbf{q} + \epsilon\mathbf{h}] - I[\mathbf{q}]}{\epsilon}.$$

We write

$$\begin{aligned} I[\mathbf{q} + \epsilon\mathbf{h}] - I[\mathbf{q}] &= \int_{t_1+\epsilon\tau_1}^{t_2+\epsilon\tau_2} L(t, \mathbf{q} + \epsilon\mathbf{h}, \dot{\mathbf{q}} + \epsilon\dot{\mathbf{h}})\, dt - I[\mathbf{q}] - \int_{t_1}^{t_2} L(t, \mathbf{q}, \dot{\mathbf{q}})\, dt \\ &= \int_{t_1}^{t_2} L(t, \mathbf{q} + \epsilon\mathbf{h}, \dot{\mathbf{q}} + \epsilon\dot{\mathbf{h}}) - L(t, \mathbf{q}, \dot{\mathbf{q}})\, dt + \int_{t_2}^{t_2+\epsilon\tau_2} L\, dt - \int_{t_1}^{t_1+\epsilon\tau_1} L\, dt. \end{aligned}$$

From (4) and the expansions

$$\int_{t_2}^{t_2+\epsilon\tau_2} L\, dt = \epsilon\tau_2 L\big|_{t_2} + o(\epsilon), \qquad \int_{t_1}^{t_1+\epsilon\tau_1} L\, dt = \epsilon\tau_1 L\big|_{t_1} + o(\epsilon)$$

$$I[\mathbf{q}+\epsilon\mathbf{h}] - I[\mathbf{q}] = \epsilon \int_{t_1}^{t_2} \sum_{i=1}^{n} \left(\frac{\partial L}{\partial q_i} - \frac{d}{dt}\frac{\partial L}{\partial \dot{q}_i} \right) h_i(t)\, dt + \epsilon \sum_{i=1}^{n} \left[\frac{\partial L}{\partial \dot{q}_i} h_i(t) \right]_{t_1}^{t_2} + \epsilon\, [L\tau]_{t_1}^{t_2} + o(\epsilon).$$

At endpoints, consistency of the perturbed curve requires

$$\mathbf{q}(t_k + \epsilon\tau_k) + \epsilon\mathbf{h}(t_k + \epsilon\tau_k) = \mathbf{q}(t_k) + \epsilon\boldsymbol{\eta}_k, \quad k = 1, 2.$$

If we expand this to first order and rearrange we find component-wise that

$$h_i(t_k) = \eta_{i,k} - \tau_k \dot{q}_i(t_k) + o(1)^6, \quad i = 1, \ldots, n,\ k = 1, 2$$

Hence, if we define, [6, p.148]

$$\text{Canonical Momentum: } p_i = \frac{\partial L}{\partial \dot{q}_i}, \qquad \text{Hamiltonian: } H = \sum_{i=1}^{n} \dot{q}_i p_i - L, \quad (5)$$

[6] Note that when we substitute this in to $I[\mathbf{q} + \epsilon\mathbf{h}] - I[\mathbf{q}]$ there is an ϵ term in front of $h_i(t_k)$ meaning we have $\epsilon o(1) = o(\epsilon)$.

we have

$$I[\mathbf{q}+\epsilon\mathbf{h}] - I[\mathbf{q}] = \epsilon \int_{t_1}^{t_2} \sum_{i=1}^{n} \left(\frac{\partial L}{\partial q_i} - \frac{d}{dt}\frac{\partial L}{\partial \dot{q}_i} \right) h_i(t)\, dt + \epsilon \left[\sum_{i=1}^{n} p_i \eta_i - H\tau \right]_{t_1}^{t_2} + o(\epsilon). \tag{6}$$

This allows us to compute the Gâteaux derivative

$$D_{\mathbf{h}} I[\mathbf{q}] = \int_{t_1}^{t_2} \sum_{i=1}^{n} \left(\frac{\partial L}{\partial q_i} - \frac{d}{dt}\frac{\partial L}{\partial \dot{q}_i} \right) h_i(t)\, dt + \left[\sum_{i=1}^{n} p_i \eta_i - H\tau \right]_{t_1}^{t_2}.$$

For $\mathbf{q}$ to be an extremal, $D_{\mathbf{h}} I[\mathbf{q}] = 0$ for all $\mathbf{h}$ which includes fixed endpoints (i.e $h_1 = h_2 = 0$) which means that $\mathbf{q}$ satisfies the Euler-Lagrange equations

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_i} - \frac{\partial L}{\partial q_i} = 0 \quad \text{for all } i = 1, \ldots, n$$

and what we will call the free-endpoint condition:

$$\left[\sum_{i=1}^{n} p_i \eta_i - H\tau \right]_{t_1}^{t_2} = 0. \tag{7}$$

3 Hamilton's Principle

Definition 3.0.1 (Mechanical Lagrangian)**.** [3, p.35] The *Lagrangian* of a mechanical system with generalised coordinates $\mathbf{q}(t) = (q_1(t), q_2(t), \ldots, q_n(t))$ is

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

where T is the kinetic energy of the system and V is the potential energy.

The Lagrangian is then a scalar function that encodes the dynamics of the system and through *Hamilton's Principle* forms an alternative formulation of classical mechanics called *Lagrangian Mechanics*, in contrast with the more familiar Newtonian formulation.

Postulate 3.0.2 (Hamilton's Principle)**.** *[3, pp.34-35] The trajectory* $\mathbf{q}(t)$ *of a system of particles from its initial position* $\mathbf{q}(t_1) = \mathbf{q}_1$ *to its final position* $\mathbf{q}(t_2) = \mathbf{q}_2$ *over the time interval* $[t_1, t_2]$ *is such that the action of the Lagrangian*

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

is extremal.

Remark 3.0.3 (Assumptions on Smoothness)**.** Of course in order to apply our Euler-Lagrange equations we must assume that $L = T - V$ has continuous second-order partial derivatives with respect to $t, q_i, \dot{q}_i$ so that the action S is Gateaux differentiable, allowing (2.3.3) to be applied.

Following these assumptions we find that the solution to the extremal problem $\mathbf{q}(t) \in S$ is described by the Euler-Lagrange equations:

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_i} - \frac{\partial L}{\partial q_i} = 0 \quad \text{for all } i = 1, \ldots, n. \tag{8}$$

In fact it is possible to recover *Newton's Second Law* by using the Euler-Lagrange equations. For further discussion on the equivalence of Newtonian and Lagrangian mechanics, [7, p.3] does the reverse inclusion of the following.

Example 3.0.4 (Recovering Newton's Second Law)**.** [6][pp.63-64]
Let $\mathbf{q}(t) = (q_1(t), q_2(t), q_3(t))$ denote the cartesian coordinates of a free particle of mass m at time t. The kinetic and potential energies are

$$T(\dot{\mathbf{q}}) = \frac{1}{2}m(\dot{q_1}^2 + \dot{q_2}^2 + \dot{q_3}^2), \qquad V(t, \mathbf{q}).$$

Hence the Lagrangian is

$$L(t, \mathbf{q}, \dot{\mathbf{q}}) = T(\dot{\mathbf{q}}) - V(t, \mathbf{q}) = \frac{1}{2}m(\dot{q_1}^2 + \dot{q_2}^2 + \dot{q_3}^2) - V(t, \mathbf{q}).$$

By the Euler-Lagrange equations (8) we find

$$m\ddot{q}_i = -\frac{\partial V}{\partial q_i}, \quad i = 1, 2, 3.$$

Defining the ith component of force on the particle as

$$f_i = -\frac{\partial V}{\partial q_i},$$

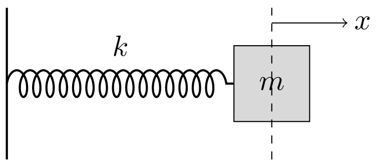
implies Newton's equation

$$\mathbf{f} = m\mathbf{a},$$

where $\mathbf{a} = \ddot{\mathbf{q}}$ is the acceleration and $\mathbf{f} = (f_1, f_2, f_3)$ is the force.

We end this section with a classical example of using our Lagrangian mechanics to formulate the equation of simple harmonic motion just as Newtonian Mechanics does.

Example 3.0.5 (Simple Harmonic Motion)**.** Consider a mass m attatched to a spring with spring constant k, moving in the x direction.



The kinetic and potential energies are

$$T = \frac{1}{2}m\dot{x}^2, \qquad V = \frac{1}{2}kx^2.$$

Hence the Lagrangian is

$$L = T - V = \frac{1}{2}m\dot{x}^2 - \frac{1}{2}kx^2.$$

Using the Euler–Lagrange equation,

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}}\right) - \frac{\partial L}{\partial x} = 0 \implies m\ddot{x} + kx = 0.$$

This is the equation of simple harmonic motion, which can be written as

$$\ddot{x} + \omega^2 x = 0, \qquad \text{where } \omega = \sqrt{\frac{k}{m}}.$$

4 Symmetries and Noether's Theorem

Having established a variational framework for mechanical systems, we now investigate transformations that leave the action invariant, known as symmetries. This section follows the methods of Van Brunt's Calculus of Variations closely, adapting proofs and generalising results where possible.

4.1 Variational Symmetries

Definition 4.1.1 (One-Parameter Variational Symmetry)**.** [6, pp.202-204] A one-parameter transformation is a smooth[7] family of mappings of the form

$$T = \theta(t, \mathbf{q}; \epsilon), \qquad Q_i = \psi_i(t, \mathbf{q}; \epsilon) \quad \text{for all } i = 1, \dots, n \tag{9}$$

where θ and ψ_i are smooth functions of $t, \mathbf{q}$ and the parameter ϵ satisfying the identity condition

$$\theta(t, \mathbf{q}; 0) = t, \qquad \psi_i(t, \mathbf{q}; 0) = q_i \quad \text{for all } i = 1, \dots, n.$$

[7] It must be at least C^2 to ensure the transformed curves and Lagrangian retain the differentiability required for the Euler-Lagrange equations.

Such a transformation is called a *variational symmetry* of the action functional

$$I[\mathbf{q}] = \int_{t_1}^{t_2} L(t, \mathbf{q}, \dot{\mathbf{q}})\, dt$$

if for every subinterval $[a, b] \subseteq [t_1, t_2]$ the action is invariant under the transformation (9), i.e.

$$\int_a^b L(t, \mathbf{q}, \dot{\mathbf{q}})\, dt = \int_{T(a)}^{T(b)} L(T, \mathbf{Q}(T), \dot{\mathbf{Q}}(T)^8)\, dT,$$

for all $\mathbf{q} \in \mathbf{C^2}[t_1, t_2]$.

Next, we state without proof a generalised version of a result from Van Brunt's Calculus of Variations, expressing the invariance of extremals under variational symmetries.

Theorem 4.1.2 (Extremal Invariance)**.** *[6, p.45] Let $L(t, \mathbf{q}, \dot{\mathbf{q}})$ be a twice continuously differentiable Lagrangian and* (9) *be a nonsingular[9] one-parameter variational symmetry.*

Then $\mathbf{q}(t)$ is an extremal of

$$I[\mathbf{q}] = \int_{t_1}^{t_2} L(t, \mathbf{q}, \dot{\mathbf{q}})\, dt$$

if and only if the transformed curve $\mathbf{Q}(T)$ is an extremal of

$$\tilde{I}[\mathbf{Q}] = \int_{T(t_1)}^{T(t_2)} L(T, \mathbf{Q}(T), \dot{\mathbf{Q}}(T))\, dT.$$

The proof is omitted due to space limitations but follows similarly to the methods in one-dimension described in [6, pp.44-46].

4.2 Noether's Theorem

Definition 4.2.1 (Conservation Law)**.** A quantity $C(t, \mathbf{q}, \dot{\mathbf{q}})$ is said to be conserved in the motion of a system if its total time derivative vanishes along any extremal $\mathbf{q}(t)$:

$$\frac{d}{dt} C(t, \mathbf{q}(t), \dot{\mathbf{q}}(t)) = 0.$$

Definition 4.2.2 (Infinitesimal Generators)**.** [6, p.208] Given a one-parameter transformation $T, Q_1(T), \ldots, Q_n(T)$, the infinitesimal generators are

$$\xi(t, \mathbf{q}) = \left.\frac{\partial \theta}{\partial \epsilon}\right|_{\epsilon=0}, \qquad \zeta_i(t, \mathbf{q}) = \left.\frac{\partial \psi_i}{\partial \epsilon}\right|_{\epsilon=0}. \tag{10}$$

[8] Note that $\dot{\mathbf{Q}}(T) = \frac{d\mathbf{Q}}{dT}$

[9] This means that the transformation is invertible and has importance within the omitted proof.

Finally we bring together what we know about free-endpoint variations in order to prove a subtle, yet beautiful theorem about the consequences of variational symmetries.

Theorem 4.2.3 (Noether's Theorem)**.** *[6, pp.210-211] Suppose a twice-continuously differentiable lagrangian* $L : [t_1, t_2]\times\mathbb{R}^n\times\mathbb{R}^n \to \mathbb{R}$*,* $L(t, \mathbf{q}(t), \dot{\mathbf{q}}(t))$*, has non-singular variational symmetries* (9) *with infinitesimal generators* ξ*,* ζ_i*.*

Then along any extremal, $\mathbf{q}(t)$*, of*

$$I[\mathbf{q}] = \int_{t_1}^{t_2} L(t, \mathbf{q}, \dot{\mathbf{q}})\, dt,$$

the quantity

$$\sum_{i=1}^{n} p_i\zeta_i - H\xi$$

is conserved.

Remark 4.2.4. The quantities p_i and H are as defined in (5).

Proof. [6, p.209] Let

$$\tilde{I}[\mathbf{q}] = \int_{a}^{b} L(t, \mathbf{q}, \dot{\mathbf{q}})\, dt,$$

where $[a, b] \subseteq [t_1, t_2]$ is an arbitrary subinterval. We regard $\tilde{I}$ as a functional with variable limits of integration.

Suppose the action is variationally invariant under the one-parameter symmetry (9)

$$T = \theta(t, \mathbf{q}; \epsilon), \qquad Q_i = \psi_i(t, \mathbf{q}; \epsilon),$$

so that

$$\tilde{I}[\mathbf{Q}] - \tilde{I}[\mathbf{q}] = \int_{T(a)}^{T(b)} L(T, \mathbf{Q}, \dot{\mathbf{Q}})\, dT - \int_{a}^{b} L(t, \mathbf{q}, \dot{\mathbf{q}})\, dt = 0.$$

Suppose that $\mathbf{q}(t)$ is an extremal. For sufficiently small ϵ, the transformed curve $\mathbf{Q}$ together with the transformed endpoints may be regarded as a variation with free endpoints. Expanding to first order gives

$$T = t + \epsilon\xi(t, \mathbf{q}) + o(\epsilon), \qquad Q_i = q_i + \epsilon\zeta_i(t, \mathbf{q}) + o(\epsilon).$$

Applying the first variation formula for free endpoints (6), we obtain

$$\tilde{I}[\mathbf{Q}] - \tilde{I}[\mathbf{q}] = \epsilon \int_{a}^{b} \sum_{i=1}^{n} \left(\frac{\partial L}{\partial q_i} - \frac{d}{dt}\frac{\partial L}{\partial \dot{q}_i} \right) h_i(t)^{10}\, dt + \epsilon \left[\sum_{i=1}^{n} p_i\zeta_i - H\xi \right]_{a}^{b} + o(\epsilon).$$

[10] Here $h_i(t)$ represents the effective variation at first order of the ith coordinate of $\mathbf{q}$ under the transformation, incorporating both the direct change in position $\mathbf{q}(t) \to \mathbf{Q}(T)$ and the contribution from the time change $t \to T$: $h_i(t) = \zeta_i(t, \mathbf{q}) - \dot{q}_i(t)\xi(t, \mathbf{q})$.

We cannot deduce the Euler–Lagrange equations from arbitrary variations vanishing at the endpoints, since the variation here is restricted. However, by the Extremal Invariance Theorem, the transformed curve $\mathbf{Q}(T)$ is also an extremal, and hence satisfies the Euler-Lagrange equations. Therefore, the integral term vanishes.

Thus,

$$\tilde{I}[\mathbf{Q}] - \tilde{I}[\mathbf{q}] = \epsilon \left[\sum_{i=1}^{n} p_i \zeta_i - H\xi \right]_a^b + o(\epsilon).$$

Since the action is invariant, the left-hand side is zero, and so

$$\left[\sum_{i=1}^{n} p_i \zeta_i - H\xi \right]_a^b = 0.$$

$[a, b]$ is arbitrary, so it follows

$$\sum_{i=1}^{n} p_i \zeta_i - H\xi$$

is constant along the extremal $\mathbf{q}(t)$. Hence,

$$\frac{d}{dt} \left(\sum_{i=1}^{n} p_i \zeta_i - H\xi \right) = 0,$$

and the quantity is conserved. $\square$

4.3 Conservation Laws

Let $\mathbf{q}(t) = (q_1(t), q_2(t), q_3(t)) \in C^2([t_1, t_2]; \mathbb{R}^3)$ describe the position of a particle (in cartesian coordinates) of mass m with kinetic and potential energies

$$T(\dot{\mathbf{q}}) = \frac{1}{2} m(\dot{q_1}^2 + \dot{q_2}^2 + \dot{q_3}^2), \qquad V(t, \mathbf{q}).$$

Using the mechanical Lagrangian (3.0.1) and Hamilton's Principle implies that the physical trajectory $\mathbf{q}$ is an extremal of

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

By Noether's Theorem we may now show that the conservation laws of energy, momentum and angular momentum then correspond to translational or rotational variational symmetries for S.

Example 4.3.1 (Conservation of Energy)**.** [6, p.212] Suppose L is variationally invariant under time translations:

$$T = t + \epsilon, \qquad Q_i = q_i \quad \text{for } i = 1, 2, 3.$$

Then the infinitesimal generators are

$$\xi = 1, \qquad \zeta_i = 0.$$

By Noether's Theorem, the conserved quantity along extremals is

$$\sum_{i=1}^{3} p_i \eta_i - H\xi = -H.$$

Hence,

$$\frac{d}{dt}(-H) = 0 \implies H = const.$$

Using (5), we compute

$$\begin{aligned} H &= \sum_{i=1}^{3} p_i \dot{q}_i - L \\ &= \sum_{i=1}^{3} \frac{\partial L}{\partial \dot{q}_i} \dot{q}_i - L \\ &= m \sum_{i=1}^{3} \dot{q}_i^2 - (T - V) \\ &= 2T - (T - V) \\ &= T + V. \end{aligned}$$

Therefore,

$$H = T + V = const$$

so invariance under time translations guarantees that total energy remains conserved along the extremal (physical) path.

Example 4.3.2 (Conservation of Linear Momentum)**.** [6, p.212] Suppose the Lagrangian is variationally invariant under translation in the q_1-direction:

$$T = t,\ Q_1 = q_1 + \epsilon,\ Q_2 = q_2,\ Q_3 = q_3.$$

Then the infinitesimal generators are

$$\xi = 0, \qquad \zeta_1 = 1, \qquad \zeta_2 = \zeta_3 = 0.$$

By Noether's Theorem, the conserved quantity along the extremal is

$$\frac{\partial L}{\partial \dot{q}_1} = const.$$

Since,

$$p_1 = \frac{\partial L}{\partial \dot{q_1}} = m\dot{q_1},$$

we have conservation of momentum in the q_1-direction:

$$p_1 = m\dot{q_1} = const$$

so invariance under translation along q_1 guarantees the linear momentum p_1 is conserved along the extremal (physical) path.

Example 4.3.3 (Conservation of Angular Momentum)**.** [6, pp.212-213] Suppose the Lagrangian is variationally invariant under rotations in the (q_1, q_2)-plane:

$$T = t, \quad Q_1 = q_1 \cos\epsilon + q_2 \sin\epsilon, \quad Q_2 = -q_1 \sin\epsilon + q_2 \cos\epsilon, \quad Q_3 = q_3$$

Expanding $\sin\epsilon$ and $\cos\epsilon$ to first order we find the infinitesimal generators are

$$\xi = 0, \qquad \zeta_1 = q_2, \qquad \zeta_2 = -q_1, \qquad \zeta_3 = 0$$

By Noether's Theorem, the conserved quantity along the extremal is

$$\sum_{i=1}^{3} \frac{\partial L}{\partial \dot{q_i}} \zeta_i = p_1 q_2 - p_2 q_1.$$

Hence,

$$\frac{d}{dt}(p_1 q_2 - p_2 q_1) = 0 \implies q_1 p_2 - q_2 p_1 = const.$$

Defining $\mathbf{p} = (p_1, p_2, p_3)$, the angular momentum is

$$\mathbf{M} = \mathbf{q} \times \mathbf{p} \implies (\mathbf{M})_3 = q_1 p_2 - q_2 p_1 = const.$$

Therefore, invariance under rotation around q_3-axis guarantees the angular momentum in the q_3 direction remains conserved along the extremal (physical) path.

5 Conclusion

As discussed in Jakob Schwichtenberg's *Physics from Symmetry* [5], these ideas extend far beyond classical mechanics. In field theory, variational principles apply to fields via Lagrangian densities, with symmetries acting on spacetime and internal degrees of freedom. In particular, the Standard Model is built on local gauge symmetries $U(1), SU(2)$, and $SU(3)$, which, through a generalised Noether's Theorem, correspond to the conservation of electric charge [5, ch.7.1.6], weak isospin [5, ch.7.7] and colour charge [5, ch.7.8.1], respectively.
What is striking is that, symmetry appears to determine, rather than merely describe the laws of physics, suggesting an unexpectedly unified picture of nature.

References

[1] Vladimir Igorevich Arnold, Karen Vogtmann, and Alan Weinstein. *Mathematical methods of classical mechanics*, volume 60. Springer, 1989.

[2] Izrail Moiseevitch Gelfand, Richard A Silverman, et al. *Calculus of variations*. Courier Corporation, 2000.

[3] Herbert Goldstein, John Safko, and Charles Poole. *Classical Mechanics*. Pearson Education UK, Harlow, third edition, 2013.

[4] B Dayanand Reddy. *Introductory functional analysis: with applications to boundary value problems and finite elements*. Number 27. Springer Science & Business Media, 1998.

[5] Jakob Schwichtenberg. *Physics from symmetry*. Springer, 2018.

[6] Bruce Van Brunt. *The calculus of variations*. Springer, 2004.

[7] Gerd Wagner and Matthew W Guthrie. Demystifying the lagrangian of classical mechanics. *arXiv preprint arXiv:1907.07069*, 2019.