1 The Calculus of Variations

Suppose we wish to examine the integral

$$I = \int_{x_1}^{x_2} F(x, y, y') \, dt \tag{1}$$

where $y(x_1) = y_1$ and $y(x_2) = y_2$. We will proceed under the assumption that the bounds of the integral and the form of the function F are given by the context of the problem. Our task is to find the curve $y = y(x)$ that makes I stationary (i.e., maximized or minimized).

Suppose $y = y(x)$ is the function that satisfies this property. Then, any variation of the function $y(x)$ will necessarily perturb (1) from its equilibrium. Consider then the following substitution:

$$y^*(x) = y(x) + \alpha \eta(x)$$

Here, $\eta(x)$ can be any function that satisfies the boundary conditions (Hence, $\eta(x_1) = \eta(x_2) = 0$), and $\alpha \in R$. Thus, by construction, we know the stationary curves of our integral I occur when $\alpha = 0$. So,

$$\left. \frac{dI}{d\alpha} \right|_{\alpha=0} = 0 \tag{2}$$

Substituting our variation into (1) and taking the relevant derivatives, we see

$$\frac{dI}{d\alpha} = \frac{d}{d\alpha} \int_{x_1}^{x_2} F(x, y^*, y'^*) \, dx$$

$$= \int_{x_1}^{x_2} \frac{\partial}{\partial \alpha} F(x, y^*, y'^*) \, dx$$

$$= \int_{x_1}^{x_2} \left[\frac{\partial F}{\partial y^*} \frac{dy^*}{d\alpha} + \frac{\partial F}{\partial y'^*} \frac{dy'^*}{d\alpha} \right] dx$$

where in the above, we have used the Leibniz Integral Rule to pass the derivative into the integrand and the chain rule to simplify. In order to satisfy our condition (2), we must have

$$\left. \frac{dI}{d\alpha} \right|_{\alpha=0} = \int_{x_1}^{x_2} \left[\frac{\partial F}{\partial y} \eta + \frac{\partial F}{\partial y'} \eta' \right] dx = 0$$

for all admissible η. Integrating the second term by parts gives the following

$$\left[\eta \frac{\partial F}{\partial y'} \right]_{x_1}^{x_2} + \int_{x_1}^{x_2} \left[\frac{\partial F}{\partial y} - \frac{d}{dx} \frac{\partial F}{\partial y'} \right] \eta \, dx = 0.$$

Notice that the first term in brackets vanishes under the boundary conditions that $\eta(x_1) = \eta(x_2) = 0$. Thus, the only way to satisfy (2) for all admissible functions η is to require

$$\frac{\partial F}{\partial y} = \frac{d}{dx} \frac{\partial F}{\partial y'} \tag{3}$$

The above is known as the Euler-Lagrange Equation. It is a fundamental underlying principle of the calculus of variations and we will soon see its significance in classical mechanics.

2 The Lagrangian and the Hamiltonian

In physics, we often talk about the *energy* within a system. Let us take a moment to define this term clearly here. Consider a system of n particles, each with mass m_i and position function $\vec{q}_i = \vec{q}_i(t)$ moving through some conservative vector field. The i'th particle's kinetic energy is

$$T_i = \frac{1}{2}m_i \dot{q}_i^2 = \frac{1}{2}m_i(\dot{x}_i^2 + \dot{y}_i^2 + \dot{z}_i^2)$$

(where the overdots represent derivatives with respect to time) and its potential energy is

$$U_i = U_i(\vec{q}_i) = U_i(x_i, y_i, z_i)$$

It will soon prove useful to consider the evolution of the difference between the total kinetic energy $T = \sum T_i$ and the total potential energy $U = \sum U_i$ of a system. Thus, we define the *action* of that system as the following functional:

$$S = \int_{t_1}^{t_2} \mathscr{L}(t, \vec{q}_i, \dot{\vec{q}}_i)\, dt \tag{4}$$

where $\mathscr{L} = T - U$ is called the *Lagrangian* of that system. This integral is in the form of (1). Thus, the stationary curves of the action must satisfy the Euler-Lagrange equation. That is,

$$\frac{\partial \mathscr{L}}{\partial \vec{q}_i} = \frac{d}{dt}\frac{\partial \mathscr{L}}{\partial \dot{\vec{q}}_i} \tag{5}$$

for each particle $i = 1, ..., n$. In this case, $\frac{\partial}{\partial \vec{q}} := [\frac{\partial}{\partial q^1}, \frac{\partial}{\partial q^2}, \frac{\partial}{\partial q^3}]$ is a vector of partial derivatives with respect to the vector components q^j written in numerator layout notation[1].

Let's consider the left side of (5):

$$\frac{\partial \mathscr{L}}{\partial \vec{q}_i} = \frac{\partial}{\partial \vec{q}_i}\{T - U\} = \frac{\partial T}{\partial \vec{q}_i} - \frac{\partial U}{\partial \vec{q}_i}$$

Because the kinetic energy of a particle does not explicitly depend on its position, $\frac{\partial T}{\partial \vec{q}_i}$ must be zero. Thus,

$$\frac{\partial \mathscr{L}}{\partial \vec{q}_i} = -\frac{\partial U}{\partial \vec{q}_i}$$

By definition, the negative of the gradient of the potential energy generated by a vector field is just the force $\vec{F}_i^{\,T}$ on the i'th particle. Thus, we see

$$\frac{\partial \mathscr{L}}{\partial \vec{q}_i} = \vec{F}_i^{\,T} \tag{6}$$

---

[1]Our convention herein will be to treat vectors, and derivatives of vectors with respect to scalars, as column vectors, and to treat derivatives of scalars with respect to vectors ac column vectors. In matrix calculus, this notation convention is referred to as *numerator layout*.

Now let's consider the right side of (5):

$$\frac{d}{dt}\frac{\partial \mathscr{L}}{\partial \dot{\vec{q}_i}} = \frac{d}{dt}\frac{\partial}{\partial \dot{\vec{q}_i}}\left\{T - U\right\} = \frac{d}{dt}\left\{\frac{\partial T}{\partial \dot{\vec{q}_i}} - \frac{\partial U}{\partial \dot{\vec{q}_i}}\right\}$$

Because the potential energy does not explicitly depend on the speed of the particle, it must be the case that $\frac{\partial U}{\partial \dot{\vec{q}_i}}$ is zero. And, $\frac{\partial T}{\partial \dot{\vec{q}_i}} = \frac{\partial}{\partial \dot{\vec{q}_i}}\{\frac{1}{2}m\dot{q}_i^2\} = m\dot{\vec{q}_i}^T$. Thus,

$$\frac{d}{dt}\frac{\partial \mathscr{L}}{\partial \dot{\vec{q}_i}} = \frac{d}{dt}\left\{m\dot{\vec{q}_i}^T\right\} = m\ddot{\vec{q}_i}^T$$

But, $\ddot{\vec{q}_i}^T$ is just the acceleration $\vec{a}_i^T$ of the particle. So,

$$\frac{d}{dt}\frac{\partial \mathscr{L}}{\partial \dot{\vec{q}_i}^T} = m\vec{a}_i^T$$

Thus, our Euler-Lagrange equation reduces to the equation $\vec{F} = m\vec{a}^T$. This is just Newton's Second Law, which is a fundamental law of physics that all systems obey. Therefore, it must be the case that the actual path an object takes upon moving from one state to another is the path that makes the action stationary.

Thus far, we have shown that Lagrange's formalism of mechanics is equivalent to Newton's. However, there exists a third mathematical formulation: Hamiltonian Mechanics. The quantity of interest here is called the *Hamiltonian*, and it is defined to be the Legendre Transform of the Lagrangian. That is,

$$\mathscr{H} = \sum_i \frac{\partial \mathscr{L}}{\partial \dot{\vec{q}_i}^T}\dot{\vec{q}_i} - \mathscr{L} \tag{7}$$

We have shown earlier that $\frac{\partial \mathscr{L}}{\partial \dot{\vec{q}_i}^T} = m\dot{\vec{q}_i}^T$, which in physics we call the momentum, $\vec{p}_i^T$, of the i'th particle. So, we can rewrite (7) in terms of that momentum as follows:

$$\mathscr{H} = \sum_i \vec{p}_i^T\dot{\vec{q}_i} - \mathscr{L} \tag{8}$$

Let us explore (8) further:

$$\begin{aligned}
\mathscr{H} &= \sum_i \vec{p}_i^T\dot{\vec{q}_i} - \mathscr{L} \\
&= \sum_i m\dot{q}_i^2 - \mathscr{L} \\
&= 2\sum_i \left(\frac{1}{2}m\dot{q}_i^2\right) - \mathscr{L} \\
&= 2T - (T - U) \\
&= T + U
\end{aligned}$$

Thus, we have shown that the Hamiltonian is just the total energy within a system. The significance of the Hamiltonian will not be addressed here, but with this notion of the Hamiltonian and our understanding of the Lagrangian, we are now ready to explore the consequences of Noether's Theorem.

3 Noether's Theorem

If a functional of the form

$$J[y] = \int_{x_1}^{x_2} F(x, y, y') \, dx$$

is invariant under the family of transformations

$$y_i^*(x) = y_i(x) + \varepsilon \mu_i(x) \qquad y_i'^*(x) = y_i'(x) + \varepsilon \mu_i'(x)$$

where ε is an infinitesimal scalar parameter, then

$$\sum_i \frac{\partial F}{\partial y_i'} \mu_i = \text{constant.} \tag{9}$$

Proof:
Because our functional is invariant under the transformation, its differential must be zero. That is,

$$dF = F(x, y_i + \varepsilon \mu_i, y_i' + \varepsilon \mu_i') - F(x, y_i, y_i') = 0$$

which, by the definition of the differential, we may write as

$$\varepsilon \sum_i \left[\frac{\partial F}{\partial y_i} \mu_i + \frac{\partial F}{\partial y_i'} \mu_i' \right] = 0$$

But, we know that the stationary curves of our functional are solutions to the Euler-Lagrange equation (3), which allows for the following substitution

$$\varepsilon \sum_i \left[\frac{d}{dx} \frac{\partial F}{\partial y_i'} \mu_i + \frac{\partial F}{\partial y_i'} \mu_i' \right] = 0$$

$$\varepsilon \frac{d}{dx} \sum_i \frac{\partial F}{\partial y_i'} \mu_i = 0$$

$$\therefore \sum_i \frac{\partial F}{\partial y_i'} \mu_i = \text{constant.}$$

∎

If we apply Noether's Theorem to the action (a functional involving the Lagrangian) of a given physical system, we will see that every continuous symmetry within the Lagrangian corresponds to a conservation law. That is, if $\mathscr{L}(t, q_i, \dot{q}_i)$ is invariant under the same family of transformations above, then

$$\sum_i \frac{\partial \mathscr{L}}{\partial \dot{q}_i} \mu_i = \text{const.}$$

Shown below are three such conservation laws.

3.1 Conservation of Linear Momentum

Suppose we have a system of n particles in some conservative vector field whose potential depends only on the relative positions of the particles. Then our Lagrangian has the form

$$\mathscr{L} = \frac{1}{2}\sum_i m_i \dot{q}_i^2 - U(\{\vec{q}_h - \vec{q}_k\}) \qquad (h = 1, ..., n-1, k > h)$$

Then it is clear that the Lagrangian is invariant under spacial translations. Such spacial translations can be represented by

$$\vec{q}_i{}^* = \vec{q}_i + \varepsilon\hat{u} \qquad \dot{\vec{q}}_i{}^* = \dot{\vec{q}}_i$$

where $\hat{u} = [u^1, u^2, u^3]^T$ is a unit vector in the direction of the translation and ε is the infinitesimal length of that translation. Then in this case we have $\mu_i = \hat{u}_i$ and thus

$$\sum_i \frac{\partial\mathscr{L}}{\partial\dot{\vec{q}}_i}\mu_i = \text{const.}$$

$$\sum_i \frac{\partial\mathscr{L}}{\partial\dot{\vec{q}}_i}\hat{u}_i = \text{const.}$$

As our convention has been to adopt numerator layout notation, we know the above multiplication is defined. Since we know that $\sum \frac{\partial\mathscr{L}}{\partial\dot{q}_i}$ is just the momentum, $\vec{p}^T$, we have

$$\sum_i \frac{\partial\mathscr{L}}{\partial\dot{\vec{q}}_i}\hat{u}_i = \text{const.}$$

$$\vec{p}^T\hat{u} = \text{const.}$$

But because $\hat{u}$ is arbitrary, it must be the case that

$$\vec{p} = \vec{const.}$$

Therefore, symmetry under spacial translation implies the conservation of linear momentum.

3.2   Conservation of Angular Momentum

Suppose again that we have a system of n particles in some conservative vector field whose potential depends now only on the distance between the particles. Then our Lagrangian has the form

$$\mathscr{L} = \frac{1}{2}\sum_i m_i \dot{q}_i^2 - U(\{|\vec{q}_h - \vec{q}_k|\}) \qquad (h = 1, ..., n-1, k > h)$$

Then it is clear that our Lagrangian is invariant under rotations about an arbitrary axis. Let $\hat{n}$ be a unit vector parallel to this axis of rotation. Then, the coordinate transformation can be described by

$$\vec{q}_i^{\,*} = \vec{q}_i + \varepsilon(\hat{n} \times \vec{q}_i) \qquad \dot{\vec{q}}_i^{T\,*} = \dot{\vec{q}}_i^{T} + \varepsilon(\hat{n} \times \dot{\vec{q}}_i^{T})$$

where ε is an infinitesimal rotation angle. Then in this case we have $\mu_i = (\hat{n} \times \vec{q}_i)$ and thus

$$\sum_{i=1}^{n} \frac{\partial \mathscr{L}}{\partial \dot{\vec{q}}_i} \mu_i = \text{const.}$$

$$\sum_{i=1}^{n} \frac{\partial \mathscr{L}}{\partial \dot{\vec{q}}_i} (\hat{n} \times \vec{q}_i) = \text{const.}$$

$$\sum_{i=1}^{n} \vec{p}_i^{\,T} (\hat{n} \times \vec{q}_i) = \text{const.}$$

$$\sum_{i=1}^{n} \hat{n}^{T} (\vec{q}_i \times \vec{p}_i) = \text{const.}$$

The vector product $\vec{q}_i \times \vec{p}_i$ is the definition of the angular momentum, $\vec{L}_i$, of the i'th particle. Thus,

$$\sum_{i=1}^{n} \hat{n}^{T} \vec{L}_i = \text{const.}$$

$$\hat{n}^{T} \vec{L} = \text{const.}$$

where L is the total angular momentum of the system. But because $\hat{n}$ is an arbitrary vector, we must conclude that

$$\vec{L} = \vec{const}.$$

Therefore, symmetry under rotation implies the conservation of angular momentum.

3.3 Conservation of Energy

Here we will use a different approach than the theorem listed above. Suppose we wish to see how the Lagrangian $\mathscr{L} = \mathscr{L}(\vec{q}_i, \dot{\vec{q}}_i^T, t)$ changes with time. Taking the total derivative, we see

$$\frac{d\mathscr{L}}{dt} = \sum_i \left[\frac{\partial \mathscr{L}}{\partial \vec{q}_i} \dot{\vec{q}}_i + \frac{\partial \mathscr{L}}{\partial \dot{\vec{q}}_i} \ddot{\vec{q}}_i \right] + \frac{\partial \mathscr{L}}{\partial t}$$

We have shown that $\mathscr{L}$ satisfies the Euler-Lagrange equation, so we can make the substitution that $\frac{\partial \mathscr{L}}{\partial \vec{q}_i} = \frac{d}{dt} \frac{\partial \mathscr{L}}{\partial \dot{\vec{q}}_i}$, which will allow us to simplify the above:

$$\frac{d\mathscr{L}}{dt} = \sum_i \left[\frac{\partial \mathscr{L}}{\partial \vec{q}_i} \dot{\vec{q}}_i + \frac{\partial \mathscr{L}}{\partial \dot{\vec{q}}_i} \ddot{\vec{q}}_i \right] + \frac{\partial \mathscr{L}}{\partial t}$$

$$\frac{d\mathscr{L}}{dt} = \sum_i \left[\frac{d}{dt} \frac{\partial \mathscr{L}}{\partial \dot{\vec{q}}_i} \dot{\vec{q}}_i + \frac{\partial \mathscr{L}}{\partial \dot{\vec{q}}_i} \ddot{\vec{q}}_i \right] + \frac{\partial \mathscr{L}}{\partial t}$$

$$\frac{d\mathscr{L}}{dt} = \frac{d}{dt} \left\{ \sum_i \frac{\partial \mathscr{L}}{\partial \dot{\vec{q}}_i} \dot{\vec{q}}_i \right\} + \frac{\partial \mathscr{L}}{\partial t}$$

$$0 = \frac{d}{dt} \left\{ \sum_i \frac{\partial \mathscr{L}}{\partial \dot{\vec{q}}_i} \dot{\vec{q}}_i - \mathscr{L} \right\} + \frac{\partial \mathscr{L}}{\partial t}$$

Recall that $\frac{\partial \mathscr{L}}{\partial \dot{\vec{q}}_i} = \vec{p}_i^T$ is the momentum of the i'th particle in our system. Imposing the simplification and explicitly writing the dot product, we have

$$\frac{d}{dt} \left\{ \sum_i \vec{p}_i^T \dot{\vec{q}}_i - \mathscr{L} \right\} + \frac{\partial \mathscr{L}}{\partial t} = 0$$

Notice that the term in brackets above is just the Hamiltonian of the system. Furthermore, because neither the kinetic nor the potential energy of most physical systems depend explicitly on time, it is reasonable to proceed under the assumption that $\frac{\partial \mathscr{L}}{\partial t} = 0$. Thus, we have

$$\frac{d\mathscr{H}}{dt} = 0 \implies \mathscr{H} = \text{constant}.$$

But, the Hamiltonian, as we have shown previously, is just the total energy within a system. Thus,

$$T + U = \text{constant}$$

Therefore, symmetry over time implies conservation of energy.

3.4 Generalizations

Thus far, we have stated Noether's Theorem as follows:

If a functional of the form

$$J[y] = \int_{x_1}^{x_2} F(x, y, y')\, dx$$

is invariant under the family of transformations

$$y_i^*(x) = y_i(x) + \varepsilon\mu_i(x) \qquad y_i'^*(x) = y_i'(x) + \varepsilon\mu_i'(x) \qquad x^* = x$$

where ε is an small scalar parameter. Then,

$$\sum_i \frac{\partial F}{\partial y_i'}\mu_i = \text{const.}$$

This is actually a special case of the more general Noether's Theorem which is presented below without proof:

Suppose we allow the following coordinate transformation:

$$x^* = \Phi(x, y, y'; \varepsilon)$$
$$y_i^* = \Psi_i(x, y, y'; \varepsilon)$$

Where $\varepsilon = 0$ corresponds to the identity transformation $x^* = x$ and $y_i^* = y_i$. Let

$$\varphi = \left.\frac{d\Phi}{d\varepsilon}\right|_{\varepsilon=0}$$
$$\psi_i = \left.\frac{d\Psi_i}{d\varepsilon}\right|_{\varepsilon=0}$$

If a functional of the form (1) is invariant under the transformations listed above, then the following is constant for any $y = y_i(x)$ that satisfies the Euler-Lagrange equation:

$$\sum_i \frac{\partial F}{\partial y_i'}\psi_i + \left(F - \sum_i \frac{\partial F}{\partial y_i'}y_i'\right)\varphi = \text{const.} \tag{10}$$

It is clear that this reduces to the theorem discussed previously if we let $\Phi(x, y, y'; \varepsilon) = x$ and $\Psi_i(x, y, y'; \varepsilon) = y_i + \varepsilon\mu_i$