The Calculus of Variations

Vasilisa Baginskaya and Stephen Zhou

August 18, 2025

Abstract

We present a short introduction to the calculus of variations. We prove the fundamental lemma and derive the Euler Lagrange equations, which we use to solve some physically inspired problems. We end with an explanation of Lagrangian mechanics and a proof of Noether's theorem.

1 Introduction

Regular calculus allows us to find the extrema of functions $f : \mathbb{R} \to \mathbb{R}$. Calculus of variations analogously lets us find the extrema of functionals, or functions of functions.

Definition 1.1 (Functional)**.** A functional is a function $I : \mathcal{F} \to \mathbb{R}$ where $\mathcal{F}$ is a set of functions.

Example 1.2. Say we want to prove that a line is the shortest path between two points. Let $y = y(x)$ be a path between (x_1, y_1) and (x_2, y_2). Then the arclength of this path is given by

$$\int_{x_1}^{x_2} \sqrt{1 + y'(x)^2}\, dx, \tag{1}$$

a functional taking y' as input.

Example 1.3. One of the first nontrivial problems to be solved using the calculus of variations was the brachistochrone problem, posed by Johann Bernoulli in 1696 and solved by Newton in 1697. The problem asks for the path between two points such that a point sliding down that path under the influence of gravity reaches the bottom point in the shortest time.

We will later show using elementary physics that the time it takes a point in a plane to slide down the curve $(x, y(x))$ from (x_1, y_2) to (x_2, y_2) is

$$\int_{x_1}^{x_2} \frac{\sqrt{1 + (y')^2}}{\sqrt{2gy}}\, dx, \tag{2}$$

where g is the force of gravity. This is a functional taking y, y' as input.

The sets of functions that we care about when minimizing are called admissible functions. The conditions we placed on admissible functions in both examples are very similar. We just required the endpoints to be two fixed points. Actually, we missed a condition: y needs to be differentiable, or else the input y' would be undefined. This int quite enough: if y' is some weird function, it might be hard to integrate when we evaluate the functional, so we will actually require that y be twice continuously differentiable, or $y \in C^2$. All other functionals in this paper will be subject to the same restrictions.

Definition 1.4. The space of admissible functions $\mathcal{A}$ consists of functions $y \in C^2[x_0, x_1]$ satisfying prescribed boundary conditions $y(x_0) = y_0$ and $y(x_1) = y_1$.

Of course, there are also other choices of admissible functions, but they tend to be harder to minimize.

Subject to the about conditions, we will find a general way to minimize functionals of the form

$$I[y] = \int_{x_0}^{x_1} F(x, y(x), y'(x))\, dx$$

where $F : [x_0, x_1] \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ is a given function called the integrand or Lagrangian. Both examples we gave are of this form. We care about functionals of this form because they are sufficient for describing many physical quantities.

Although results in the calculus of variations are much less general than in regular calculus, we will find results strong enough for us to be able to minimize action integrals.

The way we do this is to find the local extrema.

Definition 1.5 (Local Extremum)**.** A function $y_0 \in \mathcal{A}$ is a (weak) local extremum of I if there exists $\delta > 0$ such that for all $y \in \mathcal{A}$ with $\|y - y_0\|_{C^1} < \delta$, we have either $I[y] \geq I[y_0]$ (local minimum) or $I[y] \leq I[y_0]$ (local maximum).

The approach we will use to find these minima is to take one-parameter families of functions, and analyze the behavior of the functional along these families.

Example 1.6. Consider the arclength functional $\int \sqrt{1 + y'^2}\, dx$ over all $y \in C^2$ with $y(0) = 0$ and $y(1) = 1$. Then $y = x^2$ is not a local minima because $y = x^2 + \varepsilon\sqrt{(}x - x)$ has lesser arclength for sufficiently small $\varepsilon > 0$.

2 The Fundamental Lemma

The cornerstone of the calculus of variations is the fundamental lemma. This lemma lets us change an integral condition into a differential condition. The lemma itself is fairly intuitive.

Lemma 2.1 (Fundamental Lemma of Calculus of Variations)**.** Let $g : [a, b] \to \mathbb{R}$ be a continuous function. If

$$\int_a^b g(x)\eta(x)\, dx = 0$$

for all functions $\eta \in C_0^\infty(a, b) = \{\phi \in C^\infty[a, b] : \phi(a) = \phi(b) = 0\}$, then $g(x) = 0$ for all $x \in [a, b]$.

Proof. We proceed by contradiction. Suppose there exists $x_0 \in (a, b)$ such that $g(x_0) \neq 0$. Without loss of generality, assume $g(x_0) > 0$ (the case $g(x_0) < 0$ follows similarly).

By continuity of g, there exists $\delta > 0$ such that $[x_0 - \delta, x_0 + \delta] \subset (a, b)$ and

$$g(x) > \frac{g(x_0)}{2} > 0 \quad \text{for all } x \in [x_0 - \delta, x_0 + \delta]$$

$$\eta(x) = \begin{cases} \exp\left(\frac{1}{(x - x_0)^2 - \delta^2}\right) & \text{if } |x - x_0| < \delta \\ 0 & \text{otherwise} \end{cases}$$

Therefore $\eta \in C_0^\infty(a, b)$ and $\eta(x) > 0$ for $x \in (x_0 - \delta, x_0 + \delta)$.

Now we compute:

$$\int_a^b g(x)\eta(x)\, dx = \int_{x_0-\delta}^{x_0+\delta} g(x)\eta(x)\, dx \tag{3}$$

$$> \frac{g(x_0)}{2} \int_{x_0-\delta}^{x_0+\delta} \eta(x)\, dx \tag{4}$$

$$> 0 \tag{5}$$

The last inequality holds because $\eta(x) > 0$ on $(x_0 - \delta, x_0 + \delta)$ and η is not identically zero.

This contradicts our assumption that $\int_a^b g(x)\eta(x)\,dx = 0$ for all $\eta \in C_0^\infty(a,b)$. Therefore, $g(x) = 0$ for all $x \in [a,b]$. □

This lemma is more useful in it's generalized form.

Corollary 2.2 (Generalized Fundamental Lemma)**.** Let $g, h : [a,b] \to \mathbb{R}$ be continuous functions. If

$$\int_a^b [g(x)\eta(x) + h(x)\eta'(x)]\,dx = 0$$

for all $\eta \in C_0^\infty(a,b)$, then $g(x) = h'(x)$ wherever h is differentiable.

Proof. Integration by parts yields:

$$\int_a^b h(x)\eta'(x)\,dx = [h(x)\eta(x)]_a^b - \int_a^b h'(x)\eta(x)\,dx = -\int_a^b h'(x)\eta(x)\,dx$$

since $\eta(a) = \eta(b) = 0$. Therefore:

$$\int_a^b [g(x) - h'(x)]\eta(x)\,dx = 0$$

By the fundamental lemma, $g(x) - h'(x) = 0$, hence $g(x) = h'(x)$. □

3 The Euler-Lagrange Equation

We now derive our central result, which lets us minimize both functions.

Theorem 3.1 (Euler-Lagrange Equation)**.** Let $F \in C^2([x_0, x_1] \times \mathbb{R} \times \mathbb{R})$ and consider the functional

$$I[y] = \int_{x_0}^{x_1} F(x, y(x), y'(x))\,dx$$

If $y(x) \in C^2[x_0, x_1]$ is a local extremum of I subject to boundary conditions $y(x_0) = y_0$ and $y(x_1) = y_1$, then y satisfies the Euler-Lagrange equation:

$$\frac{\partial F}{\partial y} - \frac{d}{dx}\left(\frac{\partial F}{\partial y'}\right) = 0$$

Proof. Let y be a local extremum of I. For any $\eta \in C_0^\infty(x_0, x_1)$, consider the one-parameter family of functions $y_\epsilon = y + \epsilon\eta$. Note that y_ϵ satisfies the same boundary conditions as y since $\eta(x_0) = \eta(x_1) = 0$.

Define $\Phi : \mathbb{R} \to \mathbb{R}$ as $\Phi(\epsilon) = I[y_\epsilon] = \int_{x_0}^{x_1} F(x, y + \epsilon\eta, y' + \epsilon\eta')\,dx$.

Since y is a local extremum, Φ has a local extremum at $\epsilon = 0$, or $\Phi'(0) = 0$. Now we just compute $\Phi'(\varepsilon)$.

$$\begin{aligned}
\Phi'(\varepsilon) &= \frac{d}{d\varepsilon}\int_{x_0}^{x_1} F(x, y + \epsilon\eta, y' + \epsilon\eta')\,dx && (6)\\
&= \int_{x_0}^{x_1} \left[\frac{\partial F}{\partial y}\frac{\partial}{\partial \epsilon}(y + \epsilon\eta) + \frac{\partial F}{\partial y'}\frac{\partial}{\partial \epsilon}(y' + \epsilon\eta')\right] dx && (7)\\
&= \int_{x_0}^{x_1} \left[\frac{\partial F}{\partial y}\eta + \frac{\partial F}{\partial y'}\eta'\right] dx. && (8)
\end{aligned}$$

For sufficiently small ε, we are allowed to interchange differentiation and integration by the dominated convergence theorem, since $F \in C^2$. When we set $\varepsilon = 0$, we get

$$\Phi'(0) = \int_{x_0}^{x_1} \left[\frac{\partial F}{\partial y}\eta + \frac{\partial F}{\partial y'}\eta' \right] dx = 0. \tag{9}$$

So the generalized fundamental lemma implies that

$$\frac{\partial F}{\partial y} - \frac{d}{dx}\frac{\partial F}{\partial y'} = 0. \tag{10}$$

□

If the F does not depend explicitly on x (i.e., $\frac{\partial F}{\partial x} = 0$), then any solution y of the Euler Lagrange equation satisfies the conservation law:

$$\frac{d}{dx}\left(F - y'\frac{\partial F}{\partial y'}\right) = \text{constant}.$$

The expression on the left is called the Hamiltonian or the energy integral because of its use in physics to express the total energy of a system

Proof. Assume $F = F(y, y')$ and y satisfies the Euler-Lagrange equation. We compute:

$$\frac{d}{dx}\left[F - y'\frac{\partial F}{\partial y'}\right] = \frac{dF}{dx} - \frac{d}{dx}\left(y'\frac{\partial F}{\partial y'}\right) \tag{11}$$

$$= \frac{\partial F}{\partial y}y' + \frac{\partial F}{\partial y'}y'' - y''\frac{\partial F}{\partial y'} - y'\frac{d}{dx}\left(\frac{\partial F}{\partial y'}\right) \tag{12}$$

$$= y'\left[\frac{\partial F}{\partial y} - \frac{d}{dx}\left(\frac{\partial F}{\partial y'}\right)\right] \tag{13}$$

$$= 0 \tag{14}$$

The last equality follows from the Euler-Lagrange equation. Therefore, $F - y'\frac{\partial F}{\partial y'}$ is constant along any extremal. □

4 Lines and other Geodesics

As promised, we will now apply our framework to the fundamental problem of finding shortest paths. Recall that arclength is given by the integral $\int \sqrt{1 + y^{2'}}\, dx$, so we want to minimize the functional

$$\int_{x_0}^{x_1} \sqrt{1 + y'(x)^2}\, dx \tag{15}$$

over all twice differentiable functions $y(x)$.

Theorem 4.1. The shortest path between two points in the Euclidean plane $\mathbb{R}^2$ is the straight line between them.

Proof. We seek to minimize the arclength functional $L[y] = \int_{x_0}^{x_1} \sqrt{1 + (y')^2}\, dx$.

Here, $F(x, y, y') = \sqrt{1 + (y')^2}$. We compute the partial derivatives:

$$\frac{\partial F}{\partial x} = 0, \quad \frac{\partial F}{\partial y} = 0, \quad \frac{\partial F}{\partial y'} = \frac{y'}{\sqrt{1 + (y')^2}}$$

So the Euler-Lagrange equation becomes:

$$0 - \frac{d}{dx}\left(\frac{y'}{\sqrt{1+(y')^2}}\right) = 0, \tag{16}$$

which implies that $\frac{y'}{\sqrt{1+(y')^2}} = c$ for some $c \in \mathbb{R}$.

Solving for y', we get that

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

Since the right-hand side is constant, y' is a line, and the boundary conditions force it to be the unique line between (x_1, y_1) and (x_2, y_2). □

We can do the same thing for curves on any surface S parametrized as $\boldsymbol{\sigma}(u, v)$. Say we want to find the shortest path, on S between two points $\boldsymbol{\sigma}(u_1, v_1)$ and $\boldsymbol{\sigma}(u_2, v_2)$. This path is called a geodesic.

In general, the arclength of a path $\gamma(t) = \boldsymbol{\sigma}(u(t), v(t)), t \in [a, b]$ on S is given by

$$\int_a^b \sqrt{Eu'^2 + 2Fu'v' + Gv'^2}\, dx, \tag{17}$$

where $E = \boldsymbol{\sigma}_u^2, F = \boldsymbol{\sigma}_u\boldsymbol{\sigma}_v, G = \boldsymbol{\sigma}_v^2$ are the coefficients of the *first fundamental form* $E\, du^2 + F\, dudv + G\, dv^2$. Fundamental forms are a notion from differential geometry. A classic source is [1].

5 The Brachistochrone

Now we will find the brachistochrone. Recall that the brachistochrone is the curve along which a particle slides from rest under gravity in the shortest time. First we will prove the formula given in (insert) for the time it takes a particle of mass m to slide down a curve.

Theorem 5.1. The time it takes a particle of mass m to slide down a curve $y = y(x)$ from $(x_1, y(x_1))$ to $(x_2, y(x_2))$ is given by the functional

$$T[y] = \int_{x_1}^{x_2} \frac{\sqrt{1+(y')^2}}{\sqrt{2gy}}\, dx, \tag{18}$$

where g is the force of gravitational acceleration.

Proof. By energy conservation, we have

$$\frac{1}{2}mv2 = mgy$$

where v is the speed at height y. Thus $v = \sqrt{2gy}$.

The differential arclength is $ds = \sqrt{1+y'^2}dx$. The time it takes the particle to fall this length is

$$dt = \frac{ds}{v} = \frac{\sqrt{1+(y')^2}}{\sqrt{2gy}}dx. \tag{19}$$

Thus the total time is given by the functional

$$T[y] = \int_0^a \frac{\sqrt{1+(y')^2}}{\sqrt{2gy}}\, dx. \tag{20}$$

□

Theorem 5.2 (Brachistochrone Curve)**.** The brachistochrone between (x_1, y_1) and (x_2, y_2) cycloid given parametrically by:

$$x(\theta) = R(\theta - \sin\theta) \tag{21}$$
$$y(\theta) = R(1 - \cos\theta) \tag{22}$$

where R is a constant determined by the endpoint condition.

Proof. From the physical interpretation of the brachistochrone, its clear the functional has a global minimum. So we just need to use the Euler-Lagrange equations to find local minima. We have $F(x, y, y') = \frac{\sqrt{1+(y')^2}}{\sqrt{2gy}}$. Computing partial derivatives:

$$\frac{\partial F}{\partial y} = \frac{\sqrt{1+(y')^2}}{(2gy)^{3/2}} \cdot (-g) = -\frac{\sqrt{1+(y')^2}}{2y\sqrt{2gy}} = -\frac{F}{2y} \tag{23}$$

and

$$\frac{\partial F}{\partial y'} = \frac{1}{\sqrt{2gy}} \cdot \frac{y'}{\sqrt{1+(y')^2}} = \frac{y'}{\sqrt{2gy}\sqrt{1+(y')^2}} \tag{24}$$

The Euler-Lagrange equation shows that

$$-\frac{F}{2y} - \frac{d}{dx}\left(\frac{y'}{\sqrt{2gy}\sqrt{1+(y')^2}}\right) = 0 \tag{25}$$

for any local extrema y.

Since F does not depend explicitly on x, we can use the conservation law:

$$H = F - y'\frac{\partial F}{\partial y'} = \text{constant}$$

Computing:

$$H = \frac{\sqrt{1+(y')^2}}{\sqrt{2gy}} - y' \cdot \frac{y'}{\sqrt{2gy}\sqrt{1+(y')^2}} \tag{26}$$
$$= \frac{\sqrt{1+(y')^2}}{\sqrt{2gy}} - \frac{(y')^2}{\sqrt{2gy}\sqrt{1+(y')^2}} \tag{27}$$
$$= \frac{1+(y')^2-(y')^2}{\sqrt{2gy}\sqrt{1+(y')^2}} \tag{28}$$
$$= \frac{1}{\sqrt{2gy}\sqrt{1+(y')^2}} \tag{29}$$

Therefore:

$$\frac{1}{\sqrt{2gy(1+(y')^2)}} = C$$

for some constant $C > 0$. Rearranging:

$$y(1+(y')^2) = \frac{1}{2gC^2} = K$$

where $K > 0$ is a constant.

From $y(1+(y')^2) = K$, we solve for y':

$$(y')^2 = \frac{K-y}{y} = \frac{K}{y} - 1$$

For the solution to be real, we need $y \leq K$. Taking the positive square root (assuming y increases initially):

$$y' = \sqrt{\frac{K-y}{y}}$$

Separating variables:

$$\frac{dy}{dx} = \sqrt{\frac{K-y}{y}} \Rightarrow \frac{dx}{dy} = \sqrt{\frac{y}{K-y}}$$

Let $K = 2R$ for convenience and make the substitution $y = R(1-\cos\theta)$ where $\theta \in [0, \pi]$. Then:

$$dy = R\sin\theta\, d\theta$$

And:

$$\frac{y}{K-y} = \frac{R(1-\cos\theta)}{2R - R(1-\cos\theta)} = \frac{1-\cos\theta}{1+\cos\theta}$$

Using the trigonometric identity $\frac{1-\cos\theta}{1+\cos\theta} = \tan^2(\theta/2)$:

$$\sqrt{\frac{y}{K-y}} = \tan(\theta/2)$$

Therefore:

$$dx = \sqrt{\frac{y}{K-y}}\, dy = \tan(\theta/2) \cdot R\sin\theta\, d\theta$$

Using $\sin\theta = 2\sin(\theta/2)\cos(\theta/2)$:

$$dx = \tan(\theta/2) \cdot R \cdot 2\sin(\theta/2)\cos(\theta/2)\, d\theta = 2R\sin^2(\theta/2)\, d\theta$$

Using the identity $\sin^2(\theta/2) = \frac{1-\cos\theta}{2}$: $dx = R(1-\cos\theta)\, d\theta$
Integrating both expressions: $x = R\int(1-\cos\theta)\, d\theta = R(\theta - \sin\theta) + C_1$ $y = R(1-\cos\theta)$
Using the initial condition that the curve passes through $(0,0)$ when $\theta = 0$: $0 = R(0 - \sin 0) + C_1 = C_1$
Therefore, the parametric equations of the brachistochrone are:

$$x(\theta) = R(\theta - \sin\theta) \tag{30}$$
$$y(\theta) = R(1-\cos\theta) \tag{31}$$

These are the parametric equations of a cycloid - the curve traced by a fixed point on the circumference of a circle of radius R as it rolls along a straight line.

To verify this is indeed the solution, we can check that it satisfies our differential equation $y(1+(y')^2) = 2R$:
From the parametric equations: $\frac{dx}{d\theta} = R(1-\cos\theta)$, $\quad \frac{dy}{d\theta} = R\sin\theta$
Therefore: $y' = \frac{dy/d\theta}{dx/d\theta} = \frac{R\sin\theta}{R(1-\cos\theta)} = \frac{\sin\theta}{1-\cos\theta}$
Using the identity $\frac{\sin\theta}{1-\cos\theta} = \cot(\theta/2)$: $y' = \cot(\theta/2)$
Now: $(y')^2 = \cot^2(\theta/2) = \frac{\cos^2(\theta/2)}{\sin^2(\theta/2)} = \frac{1+\cos\theta}{1-\cos\theta}$
Therefore: $1+(y')^2 = 1 + \frac{1+\cos\theta}{1-\cos\theta} = \frac{1-\cos\theta+1+\cos\theta}{1-\cos\theta} = \frac{2}{1-\cos\theta}$
And: $y(1+(y')^2) = R(1-\cos\theta) \cdot \frac{2}{1-\cos\theta} = 2R$
This confirms our solution satisfies the differential equation. □

6 Noether's Theorem

One of the most beautiful theorems in the calculus of variations and physics is Noether's theorem. In physics, there are many conserved quantities, such as total energy and momentum. What Noether's Theorem tells us is that these conserved quantities are the consequence continuous symmetries of the physical world, such as time and space invariance.

To understand Noether's Theorem, we need to explain more precisely why minimizing functionals matters in physics. The reason is the *Lagrangian mechanics* formulation of physics. A mechanical system can be idealized as N point particles $(x_i(t), y_i(t), z_i(t))$ with masses m_i in space. We let the Lagrangian of this system be the kinetic energy minus the potential energy. It can be shown that this quantity depends only on the $6N$ functions $(x_i, y_i, z_i), (x_i', y_i', z_i')$ and the time t. So this definition is consistent with our previous definition of the Lagrangian, except that y can be a function $\mathbb{R}^n \to \mathbb{R}^n$ instead of $\mathbb{R} \to \mathbb{R}$. Our notation will change a little. $\mathbf{q}$ will replace y and t will replace x. The *principle of least action* tells us that the paths that are actually taken between time t_1 and t_2 are the ones that minimize the *action integral* $\int_{t_1}^{t_2} L(\mathbf{q}, \mathbf{q}', t)\, dt$. For a more detailed explanation of Lagrangian mechanics and phyics in general, see [3].

But how do we minimize these functionals? We only know how to minimize functionals of a single variable and its derivative. Fortunately, the Euler Lagrange equations generalize exactly to this more general case. In the following Theorem, the metric we use to define local extrema is $\max |\mathbf{y} - \mathbf{x}|$.

Theorem 6.1 (Euler-Lagrange Equation)**.** Let $F \in C^2([x_0, x_1] \times \mathbb{R}^n \times \mathbb{R}^n)$ and consider the functional

$$I[y] = \int_{t_t}^{x_1} F(t, \mathbf{q}(t), \mathbf{q}'(t))\, dt$$

If $\mathbf{q}(t) = (q_1(t), \ldots, q_n(t)) \in C^2$ is a local extremum of I subject to boundary conditions $\mathbf{q}(t_0) = \mathbf{q}_0$ and $\mathbf{q}(t_1) = \mathbf{q}_1$, then $\mathbf{q}$ satisfies the Euler-Lagrange equation:

$$\frac{\partial F}{\partial q_i} - \frac{d}{dt}\left(\frac{\partial F}{\partial q_i'}\right) = 0$$

The proof is effectively the exact same as the single variable case, with η being zero in all but one variable.

Theorem 6.2 (Noether's Theorem)**.** Let $L(\mathbf{q}, \mathbf{q}', t)$ be a Lagrangian. If there exists a family of transformations

$$\mathbf{q} \to \mathbf{q} + \epsilon\xi(\mathbf{q}, t), \tag{32}$$

$$t \to t + \epsilon\tau(\mathbf{q}, t) \tag{33}$$

that leaves the action invariant, then for any extremal choice of $\mathbf{q}$, the n quantites

$$J = \frac{\partial L}{\partial \mathbf{q}_i'}\xi + \left(L - \mathbf{q}'\frac{\partial L}{\partial \mathbf{q}}\right)\tau \tag{34}$$

are constant.

To prove this rigorously, we first establish the Leibniz integral rule.

Lemma 6.3 (Leibniz Integral Rule)**.** Let $f(x,t)$, $a(t)$, $b(t)$ be differentiable functions. Then

$$\frac{d}{dt}\int_{a(t)}^{b(t)} f(x,t)\, dx = f(b(t), t)\frac{db}{dt} - f(a(t), t)\frac{da}{dt} + \int_{a(t)}^{b(t)} \frac{\partial f}{\partial t}(x,t)\, dx \tag{35}$$

Proof. Using the definition of the derivative:

$$\frac{d}{dt}\int_{a(t)}^{b(t)} f(x,t)\,dx \tag{36}$$

$$= \lim_{h\to 0}\frac{1}{h}\left[\int_{a(t+h)}^{b(t+h)} f(x,t+h)\,dx - \int_{a(t)}^{b(t)} f(x,t)\,dx\right] \tag{37}$$

We can rewrite this as:

$$= \lim_{h\to 0}\frac{1}{h}\left[\int_{b(t)}^{b(t+h)} f(x,t+h)\,dx - \int_{a(t)}^{a(t+h)} f(x,t+h)\,dx + \int_{a(t)}^{b(t)} [f(x,t+h)-f(x,t)]\,dx\right] \tag{38}$$

By the dominated convergence theorem, we can exchange the limit and integral for the last term:

$$\int_{a(t)}^{b(t)} \lim_{h\to 0}\frac{f(x,t+h)-f(x,t)}{h}\,dx = \int_{a(t)}^{b(t)} \frac{\partial f}{\partial t}(x,t)\,dx \tag{39}$$

By the mean value theorem for integrals, the first two terms give:

$$f(b(t),t)\frac{db}{dt} - f(a(t),t)\frac{da}{dt} \tag{40}$$

Therefore:

$$\frac{d}{dt}\int_{a(t)}^{b(t)} f(x,t)\,dx = f(b(t),t)\frac{db}{dt} - f(a(t),t)\frac{da}{dt} + \int_{a(t)}^{b(t)} \frac{\partial f}{\partial t}(x,t)\,dx \tag{41}$$

□

Proof of Noether's Theorem. Consider the action functional:

$$S = \int_{t_1}^{t_2} L(\mathbf{q}(t),\mathbf{q}(t),t)\,dt \tag{42}$$

Under the given transformation, the action becomes:

$$S_\epsilon = \int_{t_1+\epsilon\tau(q(t_1),t_1)}^{t_2+\epsilon\tau(q(t_2),t_2)} L(q+\epsilon\xi, \dot{q}+\epsilon\dot{\xi}, t+\epsilon\tau)\,dt' \tag{43}$$

where $t' = t+\epsilon\tau$ and we need to transform the time derivative appropriately.
Since $q'(t') = q(t) + \epsilon\xi(q(t),t)$ and $t' = t + \epsilon\tau(q(t),t)$, we have:

$$\frac{dq'}{dt'} = \frac{dq'}{dt}\cdot\frac{dt}{dt'} = \frac{d}{dt}[q+\epsilon\xi]\cdot\frac{1}{1+\epsilon\dot{\tau}} = \frac{\dot{q}+\epsilon\dot{\xi}}{1+\epsilon\dot{\tau}} \tag{44}$$

To first order in ϵ:

$$\frac{dq'}{dt'} = (\dot{q}+\epsilon\dot{\xi})(1-\epsilon\dot{\tau}) = \dot{q} + \epsilon(\dot{\xi} - \dot{q}\dot{\tau}) \tag{45}$$

Now, the invariance condition states that $\frac{d}{d\epsilon}S_\epsilon|_{\epsilon=0} = 0$.
Using the Leibniz integral rule and changing back to the original time variable:

$$0 = \left.\frac{d}{d\epsilon}\right|_{\epsilon=0}\int_{t_1+\epsilon\tau_1}^{t_2+\epsilon\tau_2} L(q+\epsilon\xi, \dot{q}+\epsilon(\dot{\xi}-\dot{q}\dot{\tau}), t+\epsilon\tau)(1+\epsilon\dot{\tau})\,dt \tag{46}$$

$$= L(q(t_2),\dot{q}(t_2),t_2)\tau_2 - L(q(t_1),\dot{q}(t_1),t_1)\tau_1 \tag{47}$$

$$+ \int_{t_1}^{t_2}\left.\frac{d}{d\epsilon}\right|_{\epsilon=0} L(q+\epsilon\xi, \dot{q}+\epsilon(\dot{\xi}-\dot{q}\dot{\tau}), t+\epsilon\tau)(1+\epsilon\dot{\tau})\,dt \tag{48}$$

Computing the derivative inside the integral:

$$\begin{aligned}
&\frac{d}{d\epsilon}\bigg|_{\epsilon=0} L(q+\epsilon\xi, \dot{q}+\epsilon(\dot{\xi}-\dot{q}\dot{\tau}), t+\epsilon\tau)(1+\epsilon\dot{\tau}) && (49)\\
&= \frac{\partial L}{\partial q}\xi + \frac{\partial L}{\partial \dot{q}}(\dot{\xi}-\dot{q}\dot{\tau}) + \frac{\partial L}{\partial t}\tau + L\dot{\tau} && (50)\\
&= \frac{\partial L}{\partial q}\xi + \frac{\partial L}{\partial \dot{q}}\dot{\xi} - \frac{\partial L}{\partial \dot{q}}\dot{q}\dot{\tau} + \frac{\partial L}{\partial t}\tau + L\dot{\tau} && (51)\\
&= \frac{\partial L}{\partial q}\xi + \frac{\partial L}{\partial \dot{q}}\dot{\xi} + \left(L - \dot{q}\frac{\partial L}{\partial \dot{q}}\right)\dot{\tau} + \frac{\partial L}{\partial t}\tau && (52)
\end{aligned}$$

Therefore:

$$\begin{aligned}
0 &= L(q(t_2), \dot{q}(t_2), t_2)\tau_2 - L(q(t_1), \dot{q}(t_1), t_1)\tau_1 && (53)\\
&+ \int_{t_1}^{t_2} \left[\frac{\partial L}{\partial q}\xi + \frac{\partial L}{\partial \dot{q}}\dot{\xi} + \left(L - \dot{q}\frac{\partial L}{\partial \dot{q}}\right)\dot{\tau} + \frac{\partial L}{\partial t}\tau\right] dt && (54)
\end{aligned}$$

Using integration by parts on the second term:

$$\int_{t_1}^{t_2} \frac{\partial L}{\partial \dot{q}}\dot{\xi}\, dt = \left[\frac{\partial L}{\partial \dot{q}}\xi\right]_{t_1}^{t_2} - \int_{t_1}^{t_2} \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}}\right)\xi\, dt \tag{55}$$

For extremal paths satisfying the Euler-Lagrange equation $\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}}\right) - \frac{\partial L}{\partial q} = 0$:

$$0 = \left[L\tau + \frac{\partial L}{\partial \dot{q}}\xi\right]_{t_1}^{t_2} + \int_{t_1}^{t_2} \left[\left(L - \dot{q}\frac{\partial L}{\partial \dot{q}}\right)\dot{\tau} + \frac{\partial L}{\partial t}\tau\right] dt \tag{56}$$

Since this must hold for arbitrary endpoint times t_1, t_2, we must have:

$$\frac{d}{dt}\left[\frac{\partial L}{\partial \dot{q}}\xi + \left(L - \dot{q}\frac{\partial L}{\partial \dot{q}}\right)\tau\right] = 0 \tag{57}$$

Therefore, the Noether current

$$J = \frac{\partial L}{\partial \dot{q}}\xi + \left(L - \dot{q}\frac{\partial L}{\partial \dot{q}}\right)\tau \tag{58}$$

is conserved along extremal trajectories. □

Many conservation laws can be derived from Noether's Theorem. For example, time invariance results in the conservation of energy, translational invariance results in the conservation of momemtum, and rotational invariance results in the conservation of angular momentum.

7 Conclusion

We have only scratched the surface of the calculus of variations. There are many, many generalizations to what we have discussed. For example, we can consider functionals taking more derivatives as output, such as $F(x, y, y', y'')$, or functionals taking the derivatives of multivariable functions, such as $F(x, y, z, z_x, z_y)$, where z is a function of x and y. Our boundary conditions can also be changed. One common problem is to maximize F while another functionals G is fixed. These are collectively called isoperimetric problems, because of the case where F is area and G is arclength. The methods for solving these is similar to lagrange multipliers in regular multivariable calculus. We have also completely left out the method of finite differences, which works by approximating curves by polygons. There are also far too many applications to physics to list here. To learn more about this fascinating field, see [2] or [4] for a greater focus on applications.

References

[1] Manfredo Perdigão do Carmo. *Differential geometry of curves and surfaces. With a new preface by the author*. English. Revised and updated 2nd edition of the 1976 edition published by Prentice-Hall. Mineola, NY: Dover Publications, 2016. ISBN: 978-0-486-80699-0.

[2] I. M. Gelfand and S. V. Fomin. *Calculus of variations. Transl. from the Russian and edited by Richard A. Silverman.* English. Reprint of the 1963 original. Mineola, NY: Dover Publications, 2000. ISBN: 0-486-41448-5.

[3] Herbert Goldstein. *Classical mechanics. 2nd ed.* English. World Student Series. Addison-Wesley Series in Physics. Reading, Massachusetts, etc.: Addison-Wesley Publishing Company. XIV, 672 p. (1980). 1980.

[4] Charles R. MacCluer. *Calculus of variations. Mechanics, control, and other applications*. English. Upper Saddle River, NJ: Pearson/Prentice Hall, 2005. ISBN: 0-13-142383-5.