A Brief Introduction to the Euler-Lagrange Equation

Srikar Iyer, David Jimenez, Yizheng Li

12/15/2023

1 Historical Context

The calculus of variations deals with optimizing quantities that depend on functions. Unlike traditional calculus with derivatives, the calculus of variations seeks a function that minimizes/maximizes a certain integral. Its applications is seen in various fields including physics, engineering, economics, and biology.

A central tool in calculus of variations, the Euler-Lagrange equations are a system of second-order ODEs, whose solutions are critical points of some given action functional in the form of integrals. As the name suggests, the Euler-Lagrange equations were first derived and discovered by the Swiss and Italian mathematicians Leonhard Euler and Joseph-Louis Lagrange in the 1750s, almost three hundred years ago. Since then, Calculus of Variations has only increased in scope and applications, and is still vital regarding optimization in the modern era [3][4].

2 Prerequisites

There are a few important concepts that should be defined to understand the definition of the Euler-Lagrange equation.

Definition 2.1. (Functional) Let us consider ordinary functions as the mapping $f : \mathbb{R} \to \mathbb{R}$. Then let us define $J[x(t)]$ be the functional, which non-rigorously, takes in functions $x(t)$ as input and returns a single number. We define a functional J as the mapping $J : \mathbb{C}^\infty(\mathbb{R}) \to \mathbb{R}$. Here, $\mathbb{C}^\infty(\mathbb{R})$ represents the space of infinitely differentiable ("smooth") functions.

The most obvious, and oftentimes most convenient functional ignoring any min/max functions is the integral, since the integral on any range of an integrable function always returns a real number. So, in the following sections we will mainly consider functionals of the form $J[x] = \int_a^b f(t, x, \dot{x})dt$, and so $x \in C^1$ as requiring any higher dimensions is unnecessary. Sometimes when time t is irrelevant, the expression $J[y(x)]$ is used equivalently.

It is worth noting that the integral also denotes the "magnitude" in terms of multidimensional volume of a function, which is a number that is unique to a function and lower and upper bounds. This choice for the functional, which changes continuously as the inner function or bounds changes continuously, is helpful to understand and model variation.

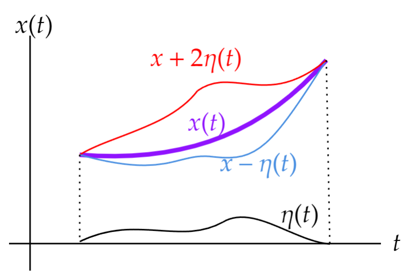
Definition 2.2. (Functional Derivative)

To find the derivative of a functional, we analyze the behavior of the functional when the original function is modified by another function, scaled by a fixed ϵ. We fix the bounds of integration to $[a, b] \subseteq \mathbb{R}$. Assuming $J[x]$ consists of finite many derivatives of x, we substitute $x(t) \to x(t) + \epsilon\eta(t)$. Then

$$\begin{aligned}
J[x+\epsilon\eta] - J[x] &= \int_a^b f(t, x+\epsilon\eta, x'+\epsilon\eta') - f(t, x, x')dt \\
&= \int_a^b \epsilon\eta\frac{\partial f}{\partial x} + \epsilon\eta\frac{d\eta}{dt}\frac{\partial f}{\partial x'} + O(\epsilon^2)dt \\
&= \left[\epsilon\eta\frac{\partial f}{\partial x}\right]_a^b + \int_a^b (\epsilon\eta(t))\left(\frac{\partial f}{\partial x} - \frac{d}{dt}\frac{\partial f}{\partial x'}\right) + O(\epsilon^2)dt
\end{aligned}$$

We call the $O(\epsilon)$ part δJ, and the *functional derivative* is defined as

$$\frac{\delta J}{\delta x} = \lim_{\epsilon\to 0}\frac{J[x(t)+\epsilon\eta(t)] - J[x(t)]}{\epsilon}$$



This expression tells us the behavior of any continuous, n-differentiable function. This leads to the functional derivative since assuming $\epsilon > 0$, calculating the derivative is similar to that of a function, except we perturb the functional by a function scaled by ϵ instead to extract the nature of the variation of the 2 functional x.

When the functional is taken at some specific points of t, we can ignore or cancel some partial components and extract a much more manageable ordinary differentiable equation. These points in question are the critical points and stationary points, both of which curves can be defined for functionals of a function.

Definition 2.3. (Stationary Points, Critical Points)

Let $x(t)$ be defined at $t = c$. Then, we have critical points wherever $f'(c) = 0$ or where $f'(c)$ is not defined.

Points where $f'(c)$ is not defined are called singular points and points where $f'(c)$ is 0 are called stationary points.

A functional J has a *stationary value (or extremum)* for input function x if the variation (or derivative) of $x = 0$.

We continue by letting δ represent the differential operator [9]. Let $\eta(a) = \eta(b) = 0$. Then, the integrated segment of the $J[x + \epsilon\eta] - J[x]$ vanishes, and the variation in x has fixed endpoints, or $\delta x(t) \equiv \epsilon\eta(t)$. If we define δJ as the $O(\epsilon)$ part of $J[x + \epsilon\eta] - J[x]$,

$$\delta J = \int_a^b \epsilon\eta(t) \left(\frac{\partial f}{\partial x} - \frac{d}{dt}\frac{\partial f}{\partial x'} \right) dt = \int_a^b \delta x(t) \left(\frac{\partial f}{\partial x} - \frac{d}{dt}\frac{\partial f}{\partial x'} \right) dt$$

The two expressions are equivalent because J and δJ are scalars, so given fixed a, b (independent of f), $\int_a^b \delta J dt = (b - a)\delta J = \delta J$ if we fix $a = 0, b = 1$. Intuitively, this doesn't affect the behavior of the functional itself is because given a function f, the functional is only scaled by integration, because multiplying a function by a constant doesn't affect the behavior of the functional at all. Hence,

$$\delta J = \delta x(t) \left(\frac{\partial f}{\partial x} - \frac{d}{dt}\frac{\partial f}{\partial x'} \right)$$

Now, we can define the functional derivative as

$$\frac{\delta J[x(t)]}{\delta x(t)} \equiv \frac{\partial f(t, x, x')}{\partial x} - \frac{d}{dt}\left(\frac{\partial f(t, x, x')}{\partial x'} \right)$$

Theorem 2.1. *(One Dimension Euler-Lagrange)*

Given a functional

$$J[x] = \int_a^b f(t, x, x')dt$$

where $x \in C^1([a, b])$ and $x' = \frac{dx}{dt}$, and that x satisfies boundary conditions $x(a) = A, x(b) = B$. If J has a stationary value at $J[x]$, then

$$\frac{\partial f}{\partial x} - \frac{d}{dt}\left(\frac{\partial f}{\partial x'} \right) = 0 \tag{1}$$

(1) is called the (one-dimensional) Euler-Lagrange equation.

The Euler-Lagrange equation is only a necessary condition but not sufficient. However in many problems, regarding physics the settings are clear that the extremum exists and is unique, so E-L is a method to find it.

3 The Euler-Lagrange Equation

3.1 Derivation of E-L

For the derivation of the one-dimensional Euler-Lagrange equation, the subsequent lemma is needed.

Lemma 3.1. *(Fundamental Lemma of the Calculus of Variation)*

Let $x(t)$ be continuous on $[a,b]$. If $\int_a^b x(t)\eta(t)dt = 0$ for all $\eta \in C^2[a,b]$ with $\eta(a) = \eta(b) = 0$, then $x(t) = 0$ for all $t \in [a,b]$.

Proof. Suppose for the sake of contradiction, that $x(t) \neq 0$ for some $c \in [a,b]$. Without loss of generality, say $x(c) > 0$. Since x is continuous, then there must be some neighborhood $c \in [c_1, c_2] \subset [a,b]$ such that $x(t) > 0$ for $t \in [c_1, c_2]$. Since x is continuous at c, then for all $\epsilon > 0$ there exists some $\delta > 0$ such that if $|c - c'| < \delta$, then $|x(c) - x(c')| < \epsilon$. Choose $\epsilon < x(c)$. If any $c' \in (c-\delta, c+\delta)$ is equal to 0, then we have $|x(c) - x(c')| > \epsilon$, which is a contradiction.

Likewise, for $c' < 0$ we get that $x(c)$ plus a positive number is greater than ϵ. So $[c_1, c_2] \subset (c-\delta, c+\delta)$. Now let $\eta(t) = (t-c_1)^2(c_2-t)^2$ for $t \in [c_1, c_2]$ and 0 otherwise. Since $\eta \in C^2$, then $\int_a^b x(t)\eta(t)dt = 0$, but due to the definition of η, this integral is equivalent to $\int_{c_1}^{c_2} x(t)\eta(t)dt$, which is greater than 0. So we have reached a contradiction and thus $x(t) = 0$ for all $t \in [a,b]$.

(Note: if $c = a$ or $c = b$, then c_1 and c_2 become a and b, respectively.) □

We now derive 1.2(1) here. The proof roughly follows [4]

Suppose a functional $J[x] = \int_a^b f(t, x(t), x'(t))dt$ satisfies $x(a) = A, x(b) = B$ with x twice differentiable on $[a,b]$. Define a perturbation function $\eta : \mathbb{R} \to \mathbb{R}$ such that $\eta \in C^1$ and $\eta(a) = \eta(b) = 0$, then for any $\epsilon \neq 0$, $\tilde{x} = x(t) + \epsilon\eta(t)$ still satisfies the boundary conditions.

Proof of Thm 1.2. Assume J obtains a stationary value at x, then when $\epsilon \to 0$, J approaches extreme value $J(x)$, hence

$$\left.\frac{d}{d\epsilon}\right|_{\epsilon=0} J[x + \epsilon\eta] = 0$$

$$\left.\frac{d}{d\epsilon}\right|_{\epsilon=0} \int_a^b f(t, x(t) + \epsilon\eta(t), x'(t) + \epsilon\eta'(t))dt = 0$$

By Leibniz's rule, we put the derivative inside

$$\int_a^b \left.\frac{\partial}{\partial\epsilon} f(t, x(t) + \epsilon\eta(t), x'(t) + \epsilon\eta'(t))\right|_{\epsilon=0} dt = 0$$

$$\int_a^b \left.\left(0 + \frac{\partial f}{\partial\tilde{x}}\frac{\partial\tilde{x}}{\partial\epsilon} + \frac{\partial f}{\partial\tilde{x}'}\frac{\partial\tilde{x}'}{\partial\epsilon}\right)\right|_{\epsilon=0} dt = 0$$

When $\epsilon = 0$, $\tilde{x} = x$. Therefore

$$\int_a^b \frac{\partial f}{\partial x}\eta + \frac{\partial f}{\partial x'}\eta' dt = 0$$

We integrate the second term inside by parts,

$$\int_a^b \frac{\partial f}{\partial x'}\eta' dt = \frac{\partial f}{\partial x'}\eta\Big|_a^b - \int_a^b \frac{d}{dt}\frac{\partial f}{\partial x'}dt = 0 - \int_a^b \frac{d}{dt}\frac{\partial f}{\partial x'}\eta dt$$

So

$$\int_a^b \left(\frac{\partial f}{\partial x} - \frac{d}{dt}\frac{\partial f}{\partial x'}\right)\eta(t) = 0$$

The value 0 is independent of the choice of $\eta(t)$, therefore by the fundamental lemma of calculus of variations, the other term must be 0. That is,

$$\frac{\partial f}{\partial x} - \frac{d}{dt}\frac{\partial f}{\partial x'} = 0$$

which is the desired equation. $\square$

3.2 Another Derivation of the E-L

We can also use the functional derivative as a starting point. We know that the functional derivative is $\frac{\partial f}{\partial x} - \frac{d}{dt}(\frac{\partial f}{\partial x'})$.

Then, let $J[x] = \int_a^b f(t, x(t), x'(t))dt$ satisfies $x(a) = A, x(b) = B$. Then, the stationary points of J are when $\delta J = \frac{dJ}{dx}\delta x = 0$. So, the necessary and sufficient condition for this to happen (since x is independent) is if $\frac{\delta J}{\delta x} = 0$.

This is equivalent to $\frac{\partial f}{\partial x} - \frac{d}{dt}(\frac{\partial f}{\partial x'}) = 0$, which is the Euler - Lagrange equation. This is also a very important interpretation of the derivation of the E-L, since it implies that the functional of any ODE finds its critical points when the ODE solves the Euler Lagrange equation, or where the functional is optimized. This implies that for any ODE, the Euler Lagrange equation returns a solution of the input ODE that, depending on the ODE, maximizes/minimizes some aspect of the flows generated from the ODE given initial conditions. [4].

3.3 Generalization to a Single Variable, Higher Order

For f with the second order derivative of x ($f(t, x, x', x'')$), make the same assumptions as above with the added conditions that $x'(a) = A'$, $x'(b) = B'$, and $\eta'(a) = \eta'(b) = 0$ with $\eta \in C^2$. As above, taking the derivative of the functional $J[x + \epsilon\eta] = \int_a^b f(t, x(t) + \epsilon\eta(t), x'(t) + \epsilon\eta(t), x''(t) + \epsilon\eta(t))dt$ becomes

$$\int_a^b \left(\frac{\partial f}{\partial t}\frac{dt}{d\epsilon} + \frac{\partial f}{\partial \tilde{x}}\frac{d\tilde{x}}{d\epsilon} + \frac{\partial f}{\partial \tilde{x}'}\frac{d\tilde{x}'}{d\epsilon} + \frac{\partial f}{\partial \tilde{x}''}\frac{d\tilde{x}''}{d\epsilon}\right) dt$$

Since t does not depend on ϵ, then when $\epsilon = 0$, we have

$$\int_a^b \left(\frac{\partial f}{\partial x}\eta(t) + \frac{\partial f}{\partial x'}\eta'(t) \right) + \frac{\partial f}{\partial x''}\eta''(t)dt = 0$$

The left part can be integrated as above. Integrating the last part by parts results in:

$$\begin{aligned}
\int_a^b \left(\frac{\partial f}{\partial x''}\eta''(t)\right)dt &= \frac{\partial f}{\partial x''}\eta'\bigg|_a^b - \int_a^b \frac{d}{dt}\frac{\partial f}{\partial x''}\eta' dt \\
&= \frac{\partial f}{\partial x''}\eta'\bigg|_a^b - \left(\frac{d}{dt}\frac{\partial f}{\partial x''}\eta\bigg|_a^b - \int_a^b \frac{d^2}{dt^2}\frac{\partial f}{\partial x''}\eta dt \right) \\
&= \left(\frac{\partial f}{\partial x''}\eta' - \frac{d}{dt}\frac{\partial f}{\partial x''}\eta \right)\bigg|_a^b + \int_a^b \frac{d^2}{dt^2}\frac{\partial f}{\partial x''}\eta dt \qquad (1)
\end{aligned}$$

Since $\eta'(a) = \eta'(b) = \eta(a) = \eta(b) = 0$, this simply becomes $\int_a^b \frac{d^2}{dt^2}\frac{\partial f}{\partial x''}\eta dt$ and combining with the integral of the left part from the previous section, we get:

$$\int_a^b \left(\frac{\partial f}{\partial x} - \frac{d}{dt}\frac{\partial f}{\partial x'} + \frac{d^2}{dt^2}\frac{\partial f}{\partial x''} \right) \eta(t)dt = 0$$

As above, the value 0 is independent of the choice of $\eta(t)$, therefore by the Fundamental Lemma of Calculus of Variations, the other term must be 0, that is

$$\frac{\partial f}{\partial x} - \frac{d}{dt}\frac{\partial f}{\partial x'} + \frac{d^2}{dt^2}\frac{\partial f}{\partial x''} = 0$$

which is the second order Euler-Lagrange equation. Note that if integration by parts had been done one more time (three times total) in (1), then an extra negative sign would have been added to the integral, thus this can be generalized for higher orders as

$$\frac{\partial f}{\partial x} - \frac{d}{dt}\frac{\partial f}{\partial x'} + \frac{d^2}{dt^2}\frac{\partial f}{\partial x''} + \cdots + (-1)^n \frac{d^n}{dt^n}\frac{\partial f}{\partial x^{(n)}} = 0$$

3.4 Generalization to Multiple Variables

Suppose $x\colon \mathbb{R} \to \mathbb{R}^n$. Then the stationary point of the action $J[x]$ can be thought of as the "intersection" of the stationary points of the actions of all of x's components. In other words, each component $x_m(t)$ of x must be a stationary point of $J[x_m]$. This results in n in different equations, namely

$$\frac{\partial f}{\partial x_m} - \frac{d}{dt}\frac{\partial f}{\partial x'_m} = 0$$

for $m = 1, 2, \ldots, n$.

Suppose x is a function of multiple variables $t_1, \ldots, t_n$. Then the action $J[x]$ becomes $\int_R f(t_1, \ldots, t_n, x, x_{t_1}, \ldots, x_{t_n})dt$, where R is some surface. As with the case of one variable, define $\eta \colon \mathbb{R}^n \to \mathbb{R}$ such that $\eta = 0$ on the boundaries of R. Then let $\tilde{x} = x(t_1, \ldots, t_n) + \epsilon\eta(t_1, \ldots, t_n)$. Similar to above, taking the derivative of the functional $J[x + \epsilon\eta]$ results in

$$
\begin{aligned}
0 &= \int_R \left(\frac{\partial f}{\partial t_1}\frac{dt_1}{d\epsilon} + \cdots + \frac{\partial f}{\partial t_n}\frac{dt_n}{d\epsilon} + \frac{\partial f}{\partial \tilde{x}}\frac{d\tilde{x}}{d\epsilon} + \frac{\partial f}{\partial \tilde{x_1}}\frac{d\tilde{x_1}}{d\epsilon} + \cdots + \frac{\partial f}{\partial \tilde{x_n}}\frac{d\tilde{x_n}}{d\epsilon} \right) d\boldsymbol{t} \\
&= \int_R \left(\frac{\partial f}{\partial \tilde{x}}\eta + \frac{\partial f}{\partial \tilde{x_1}}\eta_1 + \cdots \frac{\partial f}{\partial \tilde{x_n}}\eta_n \right) d\boldsymbol{t}
\end{aligned}
$$

since each t_n does not depend on ϵ, as above. Integrating the terms for each partial derivative by parts, as above, individually results in n different resulting $\int_{a_n}^{b_n} \frac{\partial}{\partial t_n}\frac{\partial f}{\partial \tilde{x_n}}\eta dt_n$, and putting these back into the original equation we arrive at

$$
\int_R \left(\frac{\partial f}{\partial x} + \frac{\partial}{\partial t_1}\frac{\partial f}{\partial x_1} + \cdots + \frac{\partial}{\partial t_n}\frac{\partial f}{\partial x_n} \right) \eta d\boldsymbol{t} = 0
$$

Thus, by the Fundamental Lemma of Calculus of Variations, we arrive at the Euler-Lagrange Equation of first order for multiple (n) variables:

$$
\frac{\partial f}{\partial x} + \sum_{i=1}^{n} \frac{\partial}{\partial t_i}\frac{\partial f}{\partial x_i} = 0
$$

Using a combination of these generalizations, the Euler-Lagrange equation(s) can be generalized even further, such as for multiple variables with higher order derivatives.

4 Lagrangian Mechanics

Definition 4.1. Action/Lagrangian Function

Let $J[x(t)]$ be a functional. The action is then defined as the integral $\int F(t, x, x')dt$. The function $F(t, x, x')$ will often be the difference between the kinetic and potential energy of an object during movement. In these cases, the function "L", usually called the "Lagrangian", is defined as

$$
L = T - V = \text{(Kinetic Energy)} - \text{(Potential Energy)}
$$

Theorem 4.1. *(Principle of Least Action)*

The average difference between the kinetic energy and potential energy is as small as possible for an object traveling from one point to another.

In other words, the variation of the action is 0 for the motion of an object, and thus the motion of the object is a stationary point of the action. Since this principle also coincides with the idea that objects in the real world flow and move in the path of least action, equations derived from the Euler-Lagrange equation are also known as Equations of Motion.

With the notion of "Lagrangian", many results in classical mechanics have their corresponding representations in Lagrangian mechanics. For instance, we can find the Lagrangian illustration equivalent to Newton's second Law.

Proof. We know $L = \frac{mx'^2}{2} - V(x)$. Then, $\frac{\partial L}{\partial x'} = mx'$, and $\frac{d}{dt}\frac{\partial L}{\partial x'} = mx''$. We also have $\frac{\partial L}{\partial x} = -\frac{dV}{dx}$. So, the Euler Lagrange equation states that $-\frac{dV}{dx} - mx'' = 0$, or $mx'' = -\frac{dV}{dx}$. Since $x'' = a$ and $-\frac{dV}{dx} = F$, we have $F = ma$. $\square$

ODEs or other relevant equations that model different physical scenarios can be converted to their Lagrangian form, then optimized or fact-checked with the Euler-Lagrange Equation. The following applications will be solved analytically and/or solved with the Lagrangian. Similar to the conditions of the Euler-Lagrange Equation, the Lagrangian requires that the action functional must be at a stationary point (min, max, saddle) during the entire time evolution. This, along with the Hamiltonian, form the structure for much of classical mechanics.

4.1 Lagrangian and Hamiltonian

As mentioned in class, the Hamiltonian system is of the form

$$\begin{cases} q' &= \frac{\partial H}{\partial p} \\ p' &= -\frac{\partial H}{\partial q} \end{cases}$$

Lagrangian mechanics leads to Hamiltonian mechanics naturally. While $L = T - V$, the Hamiltonian is often defined as $H = T + V$, representing total energy instead. Hamiltonian equations can always be converted to the Lagrangian with the formula

$$H = pq' - L$$

in a single dimension and the formula

$$H = \left(\sum_{i=1}^{n} q_i' \frac{\partial L}{\partial q_i'} \right) - L$$

in multiple dimensions. As a result, the Euler-Lagrange equation also works for Hamiltonian systems.

5 Some Applications

5.1 Shortest Path Between Two Points

We now prove that the shortest path between two points $x, y \in \mathbb{R}^2$ is the straight segment connecting them. This is a classical (and probably the easiest) application of the Euler-Lagrange equation.

Proof. Consider an arbitrary differentiable $y = y(x)$ passing (x_1, y_1) and (x_2, y_2). The arc length between the two points gives us a functional.

$$J[y] = \int_{x_1}^{x_2} \sqrt{1 + y'^2} dx$$

We denote $f(y) = \sqrt{1 + y'^2}$, then its partial derivatives are

$$\frac{\partial f}{\partial y} = 0 \text{ and } \frac{\partial f}{\partial y'} = \frac{y'}{\sqrt{1 + y'^2}}$$

By the Euler-Lagrange equation,

$$0 - \frac{d}{dx} \frac{y'}{\sqrt{1 + y'^2}} = 0$$

$$y' = C\sqrt{1 + y'^2}$$

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

Although trivial, this is a separable ODE. Since $y(x)$ has a constant slope and is thus a straight line. We can obtain its expression using the boundary conditions:

$$y - y_1 = \frac{y_2 - y_1}{x_2 - x_1}(x - x_1)$$

□

This purely analytical proof demonstrates an intuitive idea: if a string from point a to b held at tension (representing a line segment) is to form any other shape, more string is required to deal with the nonlinear arc. This extra string, or arc length deviation, represents the $\epsilon\eta(x)$ variation for the functional of the arc length, $J[y(x)] = \int_{x_1}^{x_2} \sqrt{1 + y'^2} dx$. and minimizing this variation as $\epsilon \to 0$ "straightens" the solution.

However, proving more complicated physical properties is more difficult analytically, so the choice of Lagrangian function and the construction of functionals will require some tricks. But before the next application, we first introduce the following property.

5.2 Beltrami identity

When f does not depend explicitly on t, we have $\frac{\partial f}{\partial t} = 0$, so the derivative of $f(t, x, x')$ is

$$\frac{df}{dt} = \frac{\partial f}{\partial t}\frac{dt}{dt} + \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial x'}\frac{dx'}{dt} = x'\frac{\partial f}{\partial x} + x''\frac{\partial f}{\partial x'}$$

Substituting $\frac{\partial f}{\partial x}$ using Euler-Lagrange, we obtain

$$\frac{df}{dt} = x'\left(\frac{d}{dt}\frac{\partial f}{\partial x'}\right) + x''\frac{\partial f}{\partial x'} = \frac{d}{dt}\left(x'\frac{\partial f}{\partial x'}\right)$$

Integrating both sides results in

$$f = x' \frac{\partial f}{\partial x'} + I \text{ , or}$$

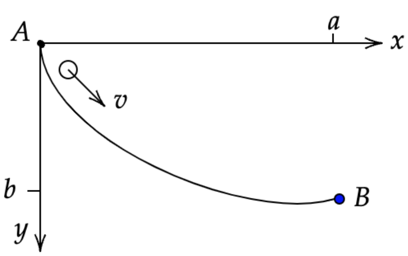
$$f - x' \frac{\partial f}{\partial x'} = I$$

where I is a constant. This is called the Beltrami identity, and its simplification of the E-L equation can oftentimes be particularly useful.

Here the quantity $I = f - x' \frac{\partial f}{\partial x'}$ is known as a *first integral* of the Euler-Lagrange equation [9].

5.3 The Brachistochrone Problem

Given two points $(0,0)$, (a,b) and $b<0$ on a vertical plane, let us find the curve of fastest descent from $(0,0)$ to (a,b) for a mass point with constant gravity, no initial velocity and no friction.



Let us try to solve this problem with a slight variant "Lagrangian" function. We take the positive direction of height to be downward, then by the conservation of total mechanical energy, $mgh = \frac{1}{2}mv^2$. At height h the speed is $\sqrt{2gh}$ with direction tangent to the curve, so a good "pseudo-Lagrangian" candidate is the time dt taken on each small arc section ds: $\frac{\sqrt{1+y'^2}}{\sqrt{2gy}}$.
Therefore the total time can be integrated as

$$T[y] = \int_0^a \frac{\sqrt{1+y'^2}}{\sqrt{2gy}} dx$$

The integrand is independent of x, so by Beltrami identity,

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

$$\sqrt{1+y'^2} - \frac{y'^2}{\sqrt{1+y'^2}} = C_1\sqrt{y}$$

$$1 + y^2 - y^2 = c\sqrt{y(1+y'^2)}$$

$$y' = \sqrt{\frac{C_2}{y} - 1}$$

This is a separable ODE, so we can solve it

$$\int \sqrt{\frac{y}{C_2 - y}} dy = \int dx$$

We apply the substitution $y = C_2 \sin^2 \theta,\ 0 \leq \theta < \pi/2$ and get

$$\int C_2 \tan \theta d C_2 \sin^2 \theta = x + C_3$$

$$C_2(\theta - \frac{1}{2} \sin 2\theta) = x + C_3$$

Therefore

$$\begin{cases} x = \frac{1}{2} C_2(2\theta - \sin 2\theta) - C_3 \\ y = \frac{1}{2} C_2(1 - 2\cos 2\theta) \end{cases}$$

The curve is a cycloid. Constants C_2 and C_3 can be obtained using the two boundary conditions.

This result can be interpreted as an approximation for the best shape of a slide that can make a marble roll down a hill, or a friction-less child slide down a playground slide the fastest. Since the result is independent of the constants, the brachistochrone can end flat or have an upward tip, which is utilized in the real world in roller coasters and skate park bowls to maximize speed.

5.4 Tautochone

This curve, also called the isochrone, represents the curve where all objects, regardless of initial position, reach the lowest point through gravity at the same time. So, we can think of this curve as a simple harmonic oscillator where all solutions reaches the bottom of a pendulum at the same time. The optimization problem that the E-L solves in this case is if the kinetic energy is maximized with respect to the arc length. Then, the kinetic component of the Lagrangian (to maximize kinetic energy) is to be optimized, or s'^2, where $s(t)$ is the arc length. Intuitively, if the arc lengths are proportional to the speed by the same constant, segments of an arc roll down a hill at the same time (since distance = speed × time. This implies the height of the curve is proportional to the arc length squared due to the formula of kinetic energy, so $y(s) = s^2$, $s^2 = x^2 + y^2$. Now,

$$dy = 2sds \Rightarrow dy^2 = 4y(dx^2 + dy^2) \Rightarrow dy^2 = 4ydx^2 + 4ydy^2$$

$$\frac{1 - 4y^2}{4y^2} dy^2 = dx^2 \Rightarrow \frac{\sqrt{1 - 4y^2}}{2y} dy = dx$$

$$\Rightarrow \int_a^b \frac{\sqrt{1 - 4y^2}}{2y} dy = \int_a^b dx$$

By the substitution $u = \sqrt{y}$,

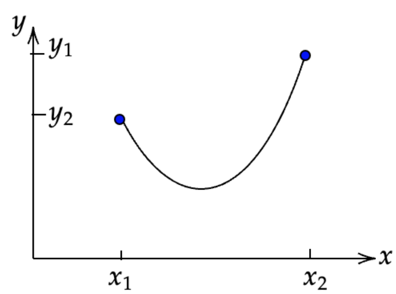
$$\begin{cases} x = \frac{1}{2}u\sqrt{1-4u^2} + \frac{1}{4}\arcsin 2u \\ y = u^2 \end{cases}$$

Using $\theta = \arcsin 2u$,

$$\begin{cases} x = \frac{1}{8}(\sin 2\theta + 2\theta) \\ y = \frac{1}{8}(1 - \cos 2\theta) \end{cases}$$

Thus, the solution to a tautochrone is also a cycloid, with differently scaled parameters compared to the Brachistochrone.

5.5 The Catenary Problem



Given two points (x_1, y_1) and (x_2, y_2), we are now interested in the shape of a homogeneous, fixed-length chain hanging between them under uniform gravity.

From physical analysis we know that the shape minimizes the potential energy. Assume that the mass per unit length ρ is 1, and that the curve is represented by $y = y(x)$ where $y(x_1) = y_1$ and $y(x_2) = y_2$, the total potential energy can be integrated along the arc length:

$$V[y] = \int_{x_1}^{x_2} y\sqrt{1 + y'^2}dx$$

Due to the restriction that the length of the chain is constant, we use the Lagrange multiplier for optimization.

$$J[y] = \int_{x_1}^{x_2} (y + \lambda)\sqrt{1 + y'^2}dx$$

The integrand $f(x, y, y') = (y + \lambda)\sqrt{1 + y'^2}$ does not explicitly depend on x, so by Beltrami identity,

$$f - y'\frac{\partial f}{\partial y'} = (y + \lambda)\sqrt{1 + y'^2} - (y + \lambda)\frac{y'^2}{\sqrt{1 + y'^2}} = C$$

$$\Rightarrow 1 + y'^2 - y'^2 = C\frac{\sqrt{1 + y'^2}}{y + \lambda}$$

$$y' = \sqrt{\left(\frac{y+\lambda}{C}\right)^2 - 1}$$

Using the substitution $\frac{y+\lambda}{C} = \cosh\theta = \frac{e^x+e^{-x}}{2}$, we have

$$\int dx = \int \frac{dy}{\sqrt{\left(\frac{y+\lambda}{C}\right)^2 - 1}} = \int \frac{C\sinh\theta}{\sinh\theta} d\theta$$

$$x + C_1 = C\theta = C\cosh^{-1}\left(\frac{y+\lambda}{C}\right)$$

Therefore

$$y = C\cosh\left(\frac{x+C_1}{C}\right) - \lambda$$

The 3 constants are determined by the two endpoints along with the length of chain.

5.6 Sturm–Liouville Problem

The E-L equation turns out to be useful in solving ODEs of some specific types, one of which is of the following form.

We want to optimize a functional I (find where $\delta I = 0$), where

$$I[y(x)] = \int_a^b [p(y')^2 + (q - \lambda p)y^2] dx$$

This functional is a bit simplified, but represents in a way kinetic energy + potential energy - a damping term (bounded by λ) which could represent friction or air resistance. The Euler-Lagrange equation, shifting variables appropriately (and setting $f = p(y')^2 + (q - \lambda p)(y^2)$), is

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

Then,

$$\frac{\partial f}{\partial y} = q' + y^2 + 2yqy' - 2\lambda ypy' - \lambda p'y^2$$

$$\frac{\partial f}{\partial y'} = 2py'y'' + p'y'$$

$$\Rightarrow q'y^2 + 2\lambda ypy' - \lambda p'y^2 - \frac{d}{dx}[2py'y'' + p'y'] = 0$$

$$\Rightarrow \frac{d}{dx}[2py'y'' + p'y'] - 2yqy' - q'y^2 = -\lambda[2ypy' + p'y^2]$$

Then, the Sturm-Liouville differential equation is a second order linear ODE of the form

$$Ly = \frac{d}{dx}\left[p(x)\frac{dy}{dx}\right] + q(x)y = -\lambda w(x)y(x)$$

for given functions $p(x), q(x), w(x)$ and linear operator L. Here, we redefine $p = 2pu'' + p', q = 2qy' - q'y$ and $w = 2py' + p'y$. The problem after this is naturally to find non-trivial eigenvalues λ's and corresponding eigen-solutions $y(x)$.

This theory is further important in applied mathematics where Sturm–Liouville problems occur very frequently, particularly when dealing with separable linear PDEs, such as the one-dimensional time-independent Schrödinger equation and wave equation. Applicable ODEs (like the Bessel equation) also exist in this format, which is also called self-adjoint. In context of functional analysis (in a Hilbert Space), $Lu = \lambda u$, and $\langle Lf, g\rangle = \langle F, Lg\rangle$. This only scratches the surface of Strum-Liouville Problems.

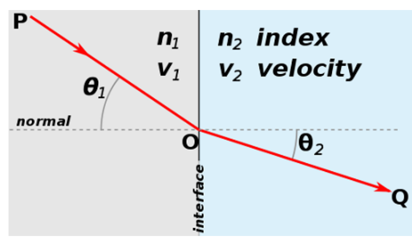
5.7 Snell's Law

When light comes in to contact with a different medium, it gets refracted, meaning it's path changes by some angle. Snell's Law is a useful tool that can be used to predict the path that light will take when it is refracted. Related to this the concept of the index of refraction, n, a property of a medium related to the speed of light in a vacuum and the speed of light in the medium. The indices of refraction for air and water, for example, are approximately 1.0002 and 1.3333, respectively.

Definition 5.1. (Index of Refraction) The index of refraction of a medium is given by

$$n = \frac{c}{v}$$

where $c \approx 3 \times 10^8$ m/s is the speed of light in a vacuum and v is the speed of light in that medium.



Theorem 5.1. *Snell's Law - The ratio of the sines of the angle of incident light and the angle of the refracted light is equivalent to the ratio of the refractive indices of the two mediums.*

$$\frac{\sin\theta_1}{\sin\theta_2} = \frac{n_2}{n_1}$$

Note that by the definition of the index of refraction $\frac{n_2}{n_1} = \frac{v_1}{v_2}$. To see why this is, we first look at Fermat's Principle:

Theorem 5.2. *(Fermat's Principle) The path that light takes is the shortest one (the one that takes the least amount of time).*

Note that this is similar to the Principle of Least Action discussed above. Now, to go into a proof of Snell's Law:

Proof. Following Fermat's Principle, we want to find a functional that takes the motion and returns the time taken. Furthermore, since the ray is a straight line (see above), we have $\int_C ds = \int \|\boldsymbol{r}'\| dt = v \int dt$. It follows, then, that the time to reach a point (in the same medium) is $\int dt = \frac{1}{v} \int_C ds$. Now note:

$$\begin{aligned}
\frac{1}{v}\int_C ds &= \frac{1}{v}\int \sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2} dt \\
&= \frac{1}{v}\int \sqrt{\left(\frac{dx}{dt}\right)^2 \left(1 + \left(\frac{dy}{dt}\cdot\frac{dt}{dx}\right)^2\right)} dt \\
&= \frac{1}{v}\int \left(\frac{dx}{dt}\right) \sqrt{1 + \left(\frac{dy}{dx}\right)^2} dt \\
&= \int \frac{1}{v}\sqrt{1 + \left(\frac{dy}{dx}\right)^2} dx
\end{aligned}$$

Now suppose the light "finishes" its path at $x = a_2$, starts at $x = a_0$, and 'switches' mediums at a_1. Then define v as $v = v_1$ for $x \in [a_0, a_1]$ and $v = v_2$ for $x \in (a_1, a_2]$. Our functional for the time is thus $J[y] = \int_{a_0}^{a_2} f(x, y(x), y'(x)dx$, where $f(x, y, y') = \frac{1}{v}\sqrt{1 + (y')^2}$. The goal is now to find the stationary point, for which we can use the Euler-Lagrange equation.

$$\begin{aligned}
\frac{\partial f}{\partial y} - \frac{d}{dx}\frac{\partial f}{\partial y'} &= 0 \\
\frac{d}{dx}\frac{\partial f}{\partial y'} &= 0 \\
\implies \frac{\partial f}{\partial y'} &= C
\end{aligned}$$

Note that $\frac{\partial f}{\partial y'} = \frac{1}{v}\frac{y'}{\sqrt{1+y'^2}}$. Now, let α_1 be the distance from a_0 to a_1, or the difference in x, and let α_2 be the difference in y from the 'start' to the moment of refraction. Note that $y' = \frac{\alpha_2}{\alpha_1}$ and $\sin\theta_1 = \frac{\alpha_2}{\sqrt{\alpha_1^2+\alpha_2^2}}$. It follows then that

$$\frac{y'}{\sqrt{1+y'^2}} = \frac{\alpha_2}{\alpha_1\sqrt{1+\frac{\alpha_2^2}{\alpha_1^2}}} = \frac{\alpha_2}{\sqrt{\alpha_1^2(1+\frac{\alpha_2^2}{\alpha_1^2})}} = \frac{\alpha_2}{\sqrt{\alpha_1^2+\alpha_2^2}} = \sin\theta_1$$

It follows now that in the first medium, $\frac{\partial f}{\partial y'} = \frac{\sin\theta_1}{v_1}$, while in the second medium $\frac{\partial f}{\partial y'} = \frac{\sin\theta_2}{v_2}$. Since $\frac{\partial f}{\partial y'}$ must constant, however, it follows $\frac{\sin\theta_1}{v_1} = \frac{\sin\theta_2}{v_2}$ and thus

$$\frac{\sin\theta_1}{\sin\theta_2} = \frac{v_1}{v_2} = \frac{n_2}{n_1}$$

□

5.8 Application to Light Diffraction and Dispersal

While the above definition uses the speed of 'general' light to define refractive index, white light is made of different colors with different wavelengths and frequencies. For example, the lower range for wavelength of red light is around 630 nm, while the higher range for violet light is around 420 nm.[5] The energy of light can be given by $E = hf$, where h is Plank's constant and f is the frequency. Conservation of energy suggests, then, that the value of the frequency must also not change in different mediums, and thus it must be the wavelength that changes. It turns out that different colors have their wavelengths changed differently, thus causing their velocities in a medium and the refractive indices of that medium to be slightly different.

Take a prism for example. While the refractive index for "white" light in crown glass is usually measured around 1.52, this is measured using a wavelength of 589 nm (near yellow or green).[8] For red and violet, the indices of refraction become 1.513 and 1.53, respectively.[2] Suppose the prism has an equilateral triangle "midsection" and you shine a light on it with a 30^o angle to the normal line. The angles θ_2 for red and violet light inside the glass would thus be around 19.297^o and 19.075^o, respectively. After some geometry, the angles of indices for the "exit" can be found to be around 40.703^o and 40.85^o, resulting in red and violet light 'emerging' from the prism at around 80 and 90 degrees respectively, thus showing why different colors emerge when shining white light on a prism. A similar phenomena can be shown with water during rain, showing why rainbows appear.

Thus the Euler-Lagrange Equation, which aids in finding Snell's Law, indirectly helps to explain various phenomena of light in addition to the other applications that we have covered.

References

[1] Fundamental Lemmas of Calculus of Variations. *Indian Institute of Science*, 2007. [Accessed Dec 2023].

[2] Refraction of Light. *Science Learning Hub*, 2020. [Accessed Dec 2023].

[3] Calculus of Variations - Wikipedia. `https://en.wikipedia.org/wiki/Calculus_of_variations`, 2023. [Online; accessed 12/15/2023].

[4] Euler–Lagrange Equation - Wikipedia. `https://en.wikipedia.org/wiki/Euler%E2%80%93Lagrange_equation`, 2023. [Online; accessed 12/15/2023].

[5] Louis A. Bloomfield. *How Things Work: the physiscs of everyday life*. John Wiley & Sons, Hoboken, NJ, 6th edition, 2016.

[6] I. Gelfand, S. Fomin, and R. Silverman. *Calculus of Variations.* Dover Publications, 2000.

[7] David Morin. The Lagrangian Method. https://scholar.harvard.edu/files/david-morin/files/cmchap6.pdf, 2007. [Accessed 2023].

[8] Leno S. Pedrotti. Basics Geometrical Optics. In *Fundamentals of Photonics*, 2008.

[9] M. Stone and P. Goldbart. *Mathematics for Physics: A Guided Tour for Graduate Students*. Cambridge University Press, 2009.

[10] Martin White. Sturm-Liouville Theory. http://w.astro.berkeley.edu/~mwhite/teachdir/sturmliouville.pdf. [Accessed Dec 2023].