The Euler-Lagrange Equation from Hamilton's "Least" Action Principle

Jacob Shapiro

December 20, 2016

1 Definition of the Lagrangian

- Convention: If $f : \mathbb{R}^m \to \mathbb{R}$ is given then $\partial_i f$ is the derivative with respect to the ith argument. If $f : \mathbb{R} \to \mathbb{R}$ then ∂f or $\dot{f}$ is the derivative (with respect to the only argument).
- Let $d \in \mathbb{N}_{\geq 1}$ be given (the number of space dimensions).
- Let $n \in \mathbb{N}_{\geq 1}$ be given (the number of particles).
- Let $V : \left(\mathbb{R}^d\right)^n \to \mathbb{R}$ (the potential) be given; We assume V is differentiable. (we treat the simplest case where the potential is time independent).
- We define $m := dn$ for breviy.
- Let $\gamma_i : \mathbb{R} \to \mathbb{R}^d$ be the trajectory of the ith particle.
- We define $\Gamma : \mathbb{R} \to \mathbb{R}^m$ as the trajectory of the whole system: $\Gamma := (\gamma_1, \ldots, \gamma_n)$.
- Thus $V \circ \Gamma : \mathbb{R} \to \mathbb{R}$.
- Given a collection of masses $(m_1, \ldots, m_n) \in (\mathbb{R}_{>0})^n$, we also define another map $T : \left(\mathbb{R}^d\right)^n \to \mathbb{R}$ via

$$T(x_1, \ldots, x_n) \quad := \quad \frac{1}{2}\sum_{i=1}^{n} m_i \|x_i\|^2 \quad \forall (x_1, \ldots, x_n) \in \left(\mathbb{R}^d\right)^n$$

  where

$$\|x_i\| \quad \equiv \quad \sqrt{(x_i)_1{}^2 + \cdots + (x_i)_d{}^2}$$

  Note that we may also consider the "flattened" version (denoted with the same letter) $T : \mathbb{R}^m \to \mathbb{R}$ with

$$(\tilde{m}_j)_{j=1}^m \quad = \quad \Big(\underbrace{m_1, \ldots, m_1}_{d \text{ times}}, \ldots, \underbrace{m_n, \ldots, m_n}_{d \text{ times}}\Big)$$

$$T(y_1, \ldots, y_m) \quad := \quad \frac{1}{2}\sum_{j=1}^{m} \tilde{m}_j y_j{}^2 \quad \forall (y_1, \ldots, y_m) \in \mathbb{R}^m$$

  Note T is differentiable as it is merely a polynomial.

- The definition of V and T was just to make things concrete but more generally we consider a Lagrangian (differentiable) map

$$L : \mathbb{R}^m \times \mathbb{R}^m \quad \to \quad \mathbb{R}$$

  and in our case

$$L\left((x_j)_{j=1}^m, (v_j)_{j=1}^m\right) \quad := \quad T\left((v_j)_{j=1}^m\right) - V\left((x_j)_{j=1}^m\right)$$

  We will also use the notation

$$L \circ \left(\Gamma, \dot{\Gamma}\right) : \underbrace{\mathbb{R}}_{\text{time}} \to \mathbb{R} \quad = \quad T \circ \dot{\Gamma} - V \circ \Gamma$$

2 Definition of the Action

- Let $(t_1,\, t_2) \in \mathbb{R}^2$ be given such that $t_1 < t_2$, and $(x_1,\, x_2) \in (\mathbb{R}^m)^2$ (the initial and final conditions).
- Define the Banach space $\mathcal{B} := C^2\left([t_1,\, t_2] \to \mathbb{R}^m\right)$ (the twice-differentiable maps from that interval to $\mathbb{R}^m$), with the supremum norm

$$\|\Gamma\|_{\mathcal{B}} \quad \equiv \quad \sup\left(\left\{\, \|\Gamma(t)\|_{\mathbb{R}^m} \mid t \in [t_1,\, t_2]\,\right\} \cup \left\{\, \left\|\dot{\Gamma}(t)\right\|_{\mathbb{R}^m} \,\middle|\, t \in [t_1,\, t_2]\,\right\}\right)$$

  Of course that norm is well-defined because a continuous image of a compact space is again compact, hence bounded.

- Define the action S_L corresponding to the Lagragian L and the time interval $[t_1,\, t_2]$ as a map $S_L : \mathcal{B} \to \mathbb{R}$ as follows

$$S_L(\Gamma) \quad := \quad \int_{t_1}^{t_2} \left(L \circ \left(\Gamma,\, \dot{\Gamma}\right)\right)(t)\, \mathrm{d}t \qquad \forall \Gamma \in \mathcal{S}$$

3 The Frechet Derivative of the Action

- Recall that the Fréchet [1] derivative of a map $S_L : \mathcal{B} \to \mathbb{R}$ between two Banach spaces at a point $\Gamma \in \mathcal{B}$ is defined as an operator $(DS_L)(\Gamma) \in \mathcal{L}(\mathcal{B},\, \mathbb{R})$ (where $\mathcal{L}(\cdot,\, \cdot)$ is the space of all continuous linear operators between two spaces) such that the following limit exists and is equal to zero:

$$\lim_{\Phi \to 0} \frac{|S_L(\Gamma + \Phi) - S_L(\Gamma) - ((DS_L)(\Gamma))(\Phi)|}{\|\Phi\|_{\mathcal{B}}} \quad = \quad 0$$

  (that is, S_L is Fréchet-differentiable at Γ if such an $(DS_L)(\Gamma)$ exists)

3.1 Claim. S_L is Fréchet differentiable on the whole of $\mathcal{B}$ and its value is

$$\begin{aligned} ((DS_L)(\Gamma))(\Phi) \quad &= \quad \sum_{j=1}^{m} \int_{t_1}^{t_2} \left(\left((\partial_j L) \circ \left(\Gamma,\, \dot{\Gamma}\right)\right)(t) - \partial\left[(\partial_{j+m} L) \circ \left(\Gamma,\, \dot{\Gamma}\right)\right](t)\right) \Phi_j(t)\, \mathrm{d}t \\ &\quad + \left[(\partial_{j+m} L)\left(\Gamma(t),\, \dot{\Gamma}(t)\right) \Phi_j(t)\right]\Big|_{t_1}^{t_2} \end{aligned}$$

Proof. Let $\varepsilon > 0$ and let $\Phi \in \mathcal{B}$ be given such that $\frac{1}{2}\varepsilon < \|\Phi\| < \varepsilon$. That implies that $\|\Phi(t)\| < \varepsilon$ and $\left\|\dot{\Phi}(t)\right\| < \varepsilon$ for all $t \in [t_1,\, t_2]$.

$$S_L(\Gamma + \Phi) \quad \equiv \quad \int_{t_1}^{t_2} \left(L \circ \left(\Gamma + \Phi,\, \dot{\Gamma} + \dot{\Phi}\right)\right)(t)\, \mathrm{d}t$$

and

$$\left(L \circ \left(\Gamma + \Phi,\, \dot{\Gamma} + \dot{\Phi}\right)\right)(t) \quad \equiv \quad L\left(\Gamma(t) + \Phi(t),\, \dot{\Gamma}(t) + \dot{\Phi}(t)\right)$$

Since $\Phi(t)$ and $\dot{\Phi}(t)$ are smaller than ε, we can make a Taylor expansion of L around $\Phi(t) = \dot{\Phi}(t) = 0$ to obtain:

$$\begin{aligned} L\left(\Gamma(t) + \Phi(t),\, \dot{\Gamma}(t) + \dot{\Phi}(t)\right) \quad &\approx \quad L\left(\Gamma(t),\, \dot{\Gamma}(t)\right) + \sum_{j=1}^{m} (\partial_j L)\left(\Gamma(t),\, \dot{\Gamma}(t)\right) \Phi_j(t) + \\ &\quad + \sum_{j=1}^{m} (\partial_{j+m} L)\left(\Gamma(t),\, \dot{\Gamma}(t)\right) \dot{\Phi}_j(t) + \mathcal{O}\left(\varepsilon^2\right) \end{aligned}$$

Next perform integration by parts on the following integral

$$\begin{aligned} \int_{t_1}^{t_2} (\partial_{j+m} L)\left(\Gamma(t),\, \dot{\Gamma}(t)\right) \dot{\Phi}_j(t)\, \mathrm{d}t \quad &= \quad \left[(\partial_{j+m} L)\left(\Gamma(t),\, \dot{\Gamma}(t)\right) \Phi_j(t)\right]\Big|_{t_1}^{t_2} - \\ &\quad - \int_{t_1}^{t_2} \left[\partial\left((\partial_{j+m} L)\left(\Gamma,\, \dot{\Gamma}\right)\right)\right](t)\, \Phi_j(t)\, \mathrm{d}t \end{aligned}$$

Hence we have (using the same (tentative) formula for $((DS_L)(\Gamma))(\Phi)$ which was introduced in the claim) to get

$$S_L(\Gamma+\Phi)-S_L(\Gamma)-((DS_L)(\Gamma))(\Phi) \quad = \quad \mathcal{O}\left(\varepsilon^2\right)(t_2-t_1)$$

We find that

$$\begin{aligned}\frac{|S_L(\Gamma+\Phi)-S_L(\Gamma)-((DS_L)(\Gamma))(\Phi)|}{\|\Phi\|_{\mathcal{B}}} &= \frac{\mathcal{O}\left(\varepsilon^2\right)|t_2-t_1|}{\|\Phi\|_{\mathcal{B}}} \\ &\leq \mathcal{O}(\varepsilon)\frac{1}{2}|t_2-t_1|\end{aligned}$$

Since $\varepsilon>0$ was arbitrary and since we can always find such Φ for a given $\varepsilon>0$, we concude the statement of the claim. □

3.2 Claim. If for some continuous $f:[t_1,\,t_2]\to\mathbb{R}$, $\int_{t_1}^{t_2} fg=0$ for all continuous $g:[t_1,\,t_2]\to\mathbb{R}$ such that $g(t_1)=g(t_2)=0$ then $f=0$. [[2] pp. 57]

Proof. Assume otherwise. Then there exists some $t_3\in[t_1,\,t_2]$ such that $|f(t_3)|>0$. Since f is continuous, there is some $\varepsilon>0$ such that $\inf(|f(B_\varepsilon(t_3))|)>c$ for some $c>0$. Pick g continuous such that $g=0$ outside $B_\varepsilon(t_3)$, $g>0$ inside $B_\varepsilon(t_3)$ and $g=1$ inside $B_{\frac{1}{2}\varepsilon}(t_3)$. Then

$$\begin{aligned}\left|\int_{t_1}^{t_2} fg\right| &\geq \left|\int_{B_\varepsilon(t_3)} fg\right| \\ &> c\left|\int_{B_\varepsilon(t_3)} g\right| \\ &> c\varepsilon\end{aligned}$$

This contradicts the fact that we should obtain zero on the left hand side. □

4 Extremum of Action Implies Euler-Lagrange Equations

4.1 Claim. The extremal points of S_L where the extremum is taken over all points such that

$$\Gamma(t_i) \quad = \quad x_i\forall \quad i\in\{\,1,\,2\,\} \tag{1}$$

is given by solutions to the (total number of m) Euler-Lagrange equations:

$$(\partial_j L)\circ\left(\Gamma,\,\dot{\Gamma}\right)-\partial\left[(\partial_{j+m}L)\circ\left(\Gamma,\,\dot{\Gamma}\right)\right] \quad = \quad 0 \quad \forall j\in\{\,1,\,\ldots,\,m\,\}$$

Proof. The extremum of a function is obtained (by definition) when its Frechet derivative is zero. That means we should seek solutions Γ to the equation

$$(DS_L)(\Gamma)|_{\mathcal{S}} \quad = \quad 0$$

where $\mathcal{S}$ is the subset of $\mathcal{B}$ such that $\Phi(t_1)=\Phi(t_2)=0$. The reason we restrict the action of the derivative to $\mathcal{S}$ is because this restriction is precisely what makes sure (1) is satisfied for every element considered for the extremum.

By the result earlier we have for all $\Phi\in\mathcal{S}$,

$$(DS_L)(\Gamma)|_{\mathcal{S}}(\Phi) \quad = \quad \sum_{j=1}^{m}\int_{t_1}^{t_2}\left(\left((\partial_j L)\circ\left(\Gamma,\,\dot{\Gamma}\right)\right)(t)-\partial\left[(\partial_{j+m}L)\circ\left(\Gamma,\,\dot{\Gamma}\right)\right](t)\right)\Phi_j(t)\,\mathrm{d}t$$

Since $(DS_L)(\Gamma)|_{\mathcal{S}}(\Phi)=0$ should hold for any $\Phi\in\mathcal{S}$, we can successively pick individual j's such that $\Phi=$

$(0, 0, \ldots, \Phi_j, \ldots, 0, 0)$ and so we actually get the m (separate) equations:

$$\int_{t_1}^{t_2} \left(\left((\partial_j L) \circ \left(\Gamma, \dot{\Gamma}\right)\right)(t) - \partial\left[(\partial_{j+m} L) \circ \left(\Gamma, \dot{\Gamma}\right)\right](t)\right) \Phi_j(t)\, \mathrm{d}t \quad = \quad 0 \quad \forall j \in \{1, \ldots, m\}$$

We now use 3.2 to conclude the statement of the claim. □

This means that the solutions to the Euler-Lagrange equations are simply the extremum points of S_L in the space of paths obeying given boundary conditions, in complete analogy to how $f'(x) \overset{!}{=} 0$ gives the extremum x of a map $f : \mathbb{R} \to \mathbb{R}$ as seen in high school.

References

[1] H. Cartan. *Differential Calculus.* Houghton Mifflin Co, 1971.

[2] MATHEMATICAL METHODS OF CLASSICAL MECHANICS. *Mathematical Methods of Classical Mechanics by V.I. Arnol'd (May 16 1989).* Springer, 1989.