

Graduate Classical Mechanics

Lecture 5: Hamilton's Principle and the Continuum Euler–Lagrange Equations

From local variations of a path to differential equations of motion

1 From the discrete action to the continuum

Lecture 4 changed the viewpoint of mechanics. Instead of only asking how a state advances from one step to the next, we fixed the endpoints of a candidate path and asked whether one scalar quantity associated with the *whole path* could select the physical trajectory.

For equally spaced times, the discrete action was

$$\mathcal{S}_d[x_0, \dots, x_N] = \sum_{n=0}^{N-1} L\left(x_n, \frac{x_{n+1} - x_n}{\Delta}\right) \Delta.$$

Varying one interior point changed only the two adjacent pieces of the action. The resulting discrete Euler–Lagrange equation was

$$\frac{\frac{\partial L_n}{\partial \dot{x}_n} - \frac{\partial L_{n-1}}{\partial \dot{x}_{n-1}}}{\Delta} - \frac{\partial L_n}{\partial x_n} = 0.$$

The first term is a discrete time derivative. As the time spacing becomes small,

$$\frac{\frac{\partial L_n}{\partial \dot{x}_n} - \frac{\partial L_{n-1}}{\partial \dot{x}_{n-1}}}{\Delta} \longrightarrow \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{x}} \right).$$

This strongly suggests the continuum equation

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{x}} - \frac{\partial L}{\partial x} = 0.$$

But suggestion is not derivation. Our goal in this lecture is to derive this equation directly from a continuous path and a continuous action.

Main structural idea

The transition is

Discrete: stationarity of a sum at every interior node → Continuum: stationarity of an integral under every local path variation.
--

The operation that turns the global integral condition into a local differential equation is *integration by parts*. It is the continuum relative of the summation-by-parts and transpose structure we met in Lecture 2.

2 A path is now a function

In the continuum, a trajectory is not a finite list of values. It is a function

$$x(t), \quad t_1 \leq t \leq t_2.$$

If we want to ask whether this path is stationary, we need a way to compare it with nearby paths.

Choose an arbitrary smooth function $\eta(t)$, and form a one-parameter family of neighboring histories

$$x_\epsilon(t) = x(t) + \epsilon\eta(t).$$

The parameter ϵ tells us how far we move away from the original path. The function $\eta(t)$ tells us the *shape* of the deformation.

If the initial and final configurations are fixed, every allowed neighboring path must share the same endpoints. Therefore

$$\eta(t_1) = \eta(t_2) = 0.$$

Inside the interval, however, $\eta(t)$ can be chosen freely.

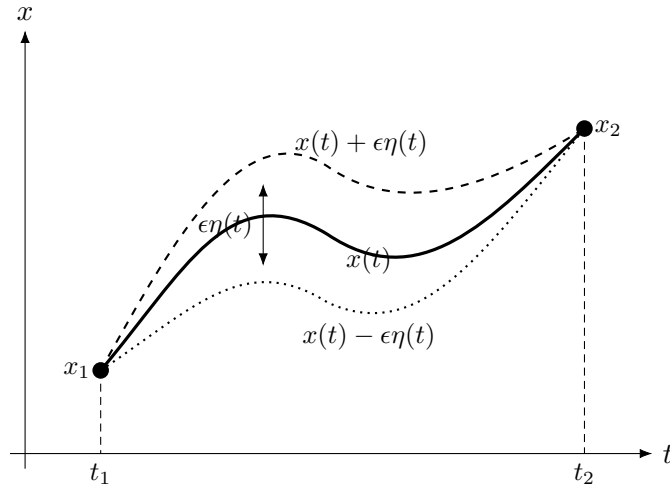


Figure 1: A path and nearby fixed-endpoint variations. The endpoints do not move, so $\eta(t_1) = \eta(t_2) = 0$, but the interior deformation is otherwise arbitrary.

2.1 Variation as a derivative in path space

For an ordinary function $V(x)$, the derivative asks how V changes when the number x is changed. For a functional $\mathcal{S}[x]$, the variation asks how $\mathcal{S}$ changes when the entire function $x(t)$ is deformed.

A precise definition is

$$\delta\mathcal{S}[x; \eta] \equiv \left. \frac{d}{d\epsilon} \mathcal{S}[x + \epsilon\eta] \right|_{\epsilon=0}.$$

Similarly,

$$\delta x(t) = \left. \frac{d}{d\epsilon} x_\epsilon(t) \right|_{\epsilon=0} = \eta(t).$$

Since differentiation with respect to t and differentiation with respect to the path parameter ϵ commute,

$$\delta\dot{x} = \left. \frac{d}{d\epsilon} \frac{dx_\epsilon}{dt} \right|_{\epsilon=0} = \frac{d}{dt}(\delta x),$$

so

$$\boxed{\delta\dot{x} = \dot{\eta}.}$$

This is the continuum version of a fact we saw discretely: moving one configuration changes the slopes attached to it.

3 Hamilton's principle

Let the action be

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

The function L is the *Lagrangian*. We allow explicit time dependence here because the variational derivation does not require time-translation symmetry. For simple conservative Cartesian mechanics,

$$L = T - V,$$

but the variational principle itself is more general than this special form.

Hamilton's principle states that the physical path is stationary under all allowed fixed-endpoint variations:

$$\boxed{\delta\mathcal{S} = 0, \quad \delta x(t_1) = \delta x(t_2) = 0.}$$

The word *stationary* is important. The action need not always be a minimum. Just as an ordinary function may have a maximum, minimum, or saddle at a point where its derivative vanishes, a functional may have different kinds of stationary paths.

4 Varying the action

Insert the neighboring path

$$x_\epsilon = x + \epsilon\eta, \quad \dot{x}_\epsilon = \dot{x} + \epsilon\dot{\eta}$$

into the action:

$$\mathcal{S}[x_\epsilon] = \int_{t_1}^{t_2} L(x + \epsilon\eta, \dot{x} + \epsilon\dot{\eta}, t) dt.$$

Differentiate with respect to ϵ and set $\epsilon = 0$. The ordinary chain rule gives

$$\boxed{\delta\mathcal{S} = \int_{t_1}^{t_2} \left[\frac{\partial L}{\partial x} \eta + \frac{\partial L}{\partial \dot{x}} \dot{\eta} \right] dt.}$$

This equation deserves to be read slowly. The path enters L in two ways:

- changing $x(t)$ changes the first argument of L , producing $\frac{\partial L}{\partial x} \eta$;
- changing $x(t)$ also changes the slope $\dot{x}(t)$, producing $\frac{\partial L}{\partial \dot{x}} \dot{\eta}$.

This is exactly the same chain-rule structure that appeared in the discrete derivation. The difficulty is that the second term contains $\dot{\eta}$, while we want a condition multiplying the arbitrary variation η itself.

5 Integration by parts moves the derivative

Consider the term

$$\int_{t_1}^{t_2} \frac{\partial L}{\partial \dot{x}} \dot{\eta} dt.$$

Apply integration by parts:

$$\int_{t_1}^{t_2} \frac{\partial L}{\partial \dot{x}} \dot{\eta} dt = \left[\frac{\partial L}{\partial \dot{x}} \eta \right]_{t_1}^{t_2} - \int_{t_1}^{t_2} \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{x}} \right) \eta dt.$$

The derivative has moved from the variation η onto the quantity $\partial L / \partial \dot{x}$. The price is a minus sign and a boundary term.

For fixed endpoints,

$$\eta(t_1) = \eta(t_2) = 0,$$

so

$$\left[\frac{\partial L}{\partial \dot{x}} \eta \right]_{t_1}^{t_2} = 0.$$

Therefore

$$\delta \mathcal{S} = \int_{t_1}^{t_2} \left[\frac{\partial L}{\partial x} - \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{x}} \right) \right] \eta(t) dt.$$

5.1 Why the boundary term matters

The boundary term vanished because we deliberately fixed the endpoint configurations. It is not meaningless algebra that should be discarded automatically. If an endpoint were free, the same term would carry information about what boundary condition the stationary path must satisfy.

This is another appearance of a recurring theme:

$$\boxed{\text{boundary terms carry physical information.}}$$

For the present derivation, fixed endpoints allow us to focus entirely on the equation in the interior.

6 From one integral to a local differential equation

Hamilton's principle requires

$$\delta \mathcal{S} = 0$$

for *every* sufficiently smooth function $\eta(t)$ that vanishes at the endpoints. We have reached

$$\int_{t_1}^{t_2} G(t)\eta(t) dt = 0,$$

where

$$G(t) = \frac{\partial L}{\partial x} - \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{x}} \right).$$

Why does this imply $G(t) = 0$ at every interior time?

Suppose instead that $G(t_*) > 0$ at some interior point t_* . If G is continuous, then it remains positive in a sufficiently small interval around t_* . We could choose a smooth variation $\eta(t)$ that is positive inside that interval and zero outside it. Then the integral would be positive, contradicting stationarity. The same argument applies if $G(t_*) < 0$.

Therefore

$$\boxed{G(t) = 0 \quad \text{for every interior } t.}$$

This statement is often called the *fundamental lemma of the calculus of variations*.

Hence

$$\frac{\partial L}{\partial x} - \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{x}} \right) = 0,$$

or, in the conventional order,

$$\boxed{\frac{d}{dt} \frac{\partial L}{\partial \dot{x}} - \frac{\partial L}{\partial x} = 0.}$$

This is the *Euler–Lagrange equation*.

A global statement about the action becomes a local differential equation because the interior variation can be localized arbitrarily close to any time.

7 The discrete and continuum structures side by side

The continuum derivation is not a new trick unrelated to Lecture 4. It is the limiting version of exactly the same structure.

Discrete path	Continuous path
Values x_n at nodes	Function $x(t)$
Variation δx_n at an interior node	Variation $\eta(t) = \delta x(t)$ in the interior
Neighboring slope differences	Derivative $\dot{x}$
$\delta \dot{x}_n$ comes from neighboring node changes	$\delta \dot{x} = d(\delta x)/dt$
Sum over steps	Integral over time
Summation by parts / transpose	Integration by parts / adjoint
Equation at every interior node	Differential equation at every interior time

There is also a useful operator statement. Let

$$A_t = \frac{d}{dt}.$$

Then, with the time inner product

$$\langle f, g \rangle = \int_{t_1}^{t_2} f(t)g(t) dt,$$

integration by parts says

$$\langle f, A_t g \rangle = [fg]_{t_1}^{t_2} + \langle -A_t f, g \rangle.$$

Thus the formal adjoint is

$$\boxed{A_t^T = -\frac{d}{dt}}$$

when the boundary term is handled separately.

The variation can therefore be read schematically as

$$\delta\mathcal{S} = \langle L_x, \delta x \rangle + \langle L_{\dot{x}}, A_t \delta x \rangle,$$

and after moving the derivative,

$$\delta\mathcal{S} = \langle L_x + A_t^T L_{\dot{x}}, \delta x \rangle + \text{boundary term}.$$

Stationarity of the interior therefore gives

$$L_x + A_t^T L_{\dot{x}} = 0,$$

which is simply the Euler–Lagrange equation again.

This is the continuum relative of the A^T structure in Lecture 2: a quantity associated with differences is moved back to the nodes—or here, back to the underlying function—by the adjoint operator.

8 Example 1: the free particle

Take

$$\boxed{L = \frac{1}{2}m\dot{x}^2}.$$

Then

$$\frac{\partial L}{\partial \dot{x}} = m\dot{x}, \quad \frac{\partial L}{\partial x} = 0.$$

Euler–Lagrange gives

$$\frac{d}{dt}(m\dot{x}) = 0.$$

For constant mass,

$$\boxed{m\ddot{x} = 0}.$$

Thus

$$\dot{x} = \text{constant}, \quad x(t) = x(t_0) + v_0(t - t_0).$$

We have now met the same free motion from three viewpoints:

$$\boxed{K = 0 \iff p = \text{constant} \iff \delta\mathcal{S} = 0 \text{ for } L = \frac{1}{2}m\dot{x}^2}.$$

The physical content has not changed; the organization of the problem has.

9 Example 2: a conservative particle

For a particle in a position-dependent potential,

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

Then

$$\frac{\partial L}{\partial \dot{x}} = m\dot{x}, \quad \frac{\partial L}{\partial x} = -V'(x).$$

Euler–Lagrange becomes

$$\frac{d}{dt}(m\dot{x}) + V'(x) = 0.$$

For constant mass,

$$m\ddot{x} = -V'(x) = F(x).$$

Thus Hamilton’s principle reproduces Newton’s equation for conservative mechanics.

9.1 Generalized momentum

The quantity

$$p \equiv \frac{\partial L}{\partial \dot{x}}$$

is called the *generalized momentum* conjugate to x . In the simple Cartesian example above,

$$p = m\dot{x},$$

so it agrees with ordinary mechanical momentum.

The definition is more important than this special coincidence. When we move to generalized coordinates, the conjugate momentum need not look like $m\dot{q}$, but the quantity $\partial L/\partial \dot{q}$ continues to play the same structural role.

With this definition, the Euler–Lagrange equation can be written

$$\dot{p} = \frac{\partial L}{\partial x}.$$

For $L = T - V$, the right-hand side is the conservative force.

10 A variational problem outside mechanics: the shortest path

The Euler–Lagrange equation is not fundamentally a law of mechanics. It is a mathematical statement about stationary functionals. Mechanics is one especially important application.

Consider a plane curve written as $x = x(y)$. Its arc length between y_1 and y_2 is

$$\ell[x] = \int_{y_1}^{y_2} \sqrt{1 + \left(\frac{dx}{dy}\right)^2} dy.$$

Define

$$x' \equiv \frac{dx}{dy}, \quad f(x, x', y) = \sqrt{1 + (x')^2}.$$

The stationary curve satisfies the Euler–Lagrange equation with y playing the role that time played in mechanics:

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

Because f has no explicit dependence on x ,

$$\frac{\partial f}{\partial x} = 0.$$

Also,

$$\frac{\partial f}{\partial x'} = \frac{x'}{\sqrt{1 + (x')^2}}.$$

Therefore

$$\boxed{\frac{d}{dy} \left(\frac{x'}{\sqrt{1 + (x')^2}} \right) = 0.}$$

The quantity inside parentheses is constant. Since the function

$$z \mapsto \frac{z}{\sqrt{1 + z^2}}$$

is one-to-one, x' itself is constant. Hence

$$\boxed{x(y) = ay + b.}$$

The stationary path between two points in Euclidean space is the straight line; for this problem the stationary path is indeed the minimum-length path.

10.1 Fermat's principle as the same structure

If light moves through a medium with local speed $u(x, y)$, the travel time along a path $x(y)$ is

$$T[x] = \int_{y_1}^{y_2} \frac{\sqrt{1 + (x')^2}}{u(x, y)} dy.$$

Fermat's principle states that the physical ray makes this travel time stationary under nearby fixed-endpoint variations:

$$\boxed{\delta T = 0.}$$

If the refractive index is $n = c_0/u$, the same functional is proportional to

$$\int n(x, y) \sqrt{1 + (x')^2} dy.$$

The mechanics action

$$\int L(x, \dot{x}, t) dt$$

is one member of this broader family of variational problems.

11 Many coordinates

Real mechanical systems usually have more than one coordinate. Let

$$\mathbf{q}(t) = (q_1(t), \dots, q_N(t))$$

and let

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

Choose an arbitrary vector variation

$$\delta \mathbf{q}(t) = \boldsymbol{\eta}(t) = (\eta_1(t), \dots, \eta_N(t)),$$

with

$$\boldsymbol{\eta}(t_1) = \boldsymbol{\eta}(t_2) = \mathbf{0}.$$

The first variation is

$$\delta \mathcal{S} = \sum_{i=1}^N \int_{t_1}^{t_2} \left[\frac{\partial L}{\partial q_i} \eta_i + \frac{\partial L}{\partial \dot{q}_i} \dot{\eta}_i \right] dt.$$

Integrating the second term by parts for each coordinate gives

$$\delta \mathcal{S} = \sum_{i=1}^N \int_{t_1}^{t_2} \left[\frac{\partial L}{\partial q_i} - \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_i} \right) \right] \eta_i dt.$$

The functions $\eta_i(t)$ can be varied independently. Therefore each coefficient must vanish:

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{q}_i} - \frac{\partial L}{\partial q_i} = 0, \quad i = 1, \dots, N.$$

In vector notation,

$$\frac{d}{dt} \nabla_{\dot{\mathbf{q}}} L - \nabla_{\mathbf{q}} L = \mathbf{0}.$$

The conjugate momenta are

$$p_i = \frac{\partial L}{\partial \dot{q}_i}, \quad \mathbf{p} = \nabla_{\dot{\mathbf{q}}} L.$$

12 Why generalized coordinates are natural

The notation q_i is deliberate. Nothing in the derivation required the coordinates to be Cartesian positions.

Suppose, for example, that a particle is constrained to a circle of radius R :

$$x^2 + y^2 = R^2.$$

If we insist on using x and y , then the coordinates are not independent and Newton's equations require a constraint force to keep the particle on the circle.

But the allowed configurations can be described by one coordinate,

$$\boxed{x = R \cos \theta, \quad y = R \sin \theta.}$$

The single number θ labels every allowed configuration. Its time derivative determines the velocity along the circle:

$$\dot{x}^2 + \dot{y}^2 = R^2 \dot{\theta}^2.$$

Thus the kinetic energy becomes

$$T = \frac{1}{2} m R^2 \dot{\theta}^2.$$

The variational formalism is already prepared to work directly with θ as the coordinate.

This is the central reason generalized coordinates are powerful: instead of describing motion in a large coordinate space and then enforcing constraints, we can often choose coordinates that describe only the configurations that are actually allowed.

We will develop this systematically next.

13 Structural note: adding a total derivative

There is an important freedom in the choice of Lagrangian. Suppose

$$L'(q, \dot{q}, t) = L(q, \dot{q}, t) + \frac{d}{dt} F(q, t).$$

Then the new action is

$$\begin{aligned} \mathcal{S}' &= \int_{t_1}^{t_2} L' dt \\ &= \mathcal{S} + \int_{t_1}^{t_2} \frac{dF}{dt} dt \\ &= \mathcal{S} + F(q(t_2), t_2) - F(q(t_1), t_1). \end{aligned}$$

For fixed endpoints, the additional term depends only on the fixed boundary data, so its variation is zero:

$$\delta \mathcal{S}' = \delta \mathcal{S}.$$

Therefore

$$\boxed{L \quad \text{and} \quad L + \frac{dF}{dt} \quad \text{produce the same Euler–Lagrange equations.}}$$

The Lagrangian is therefore not unique. Once again, the reason is visible in the boundary term.

This freedom will become useful later, especially when we discuss symmetries and more general interactions.

14 Summary

The central derivation of this lecture is

$$\boxed{\mathcal{S}[x] = \int_{t_1}^{t_2} L(x, \dot{x}, t) dt, \quad \delta \mathcal{S} = 0.}$$

A path variation

$$x \rightarrow x + \epsilon \eta$$

changes both the position and the slope:

$$\delta x = \eta, \quad \delta \dot{x} = \dot{\eta}.$$

The chain rule gives

$$\delta \mathcal{S} = \int (L_x \eta + L_{\dot{x}} \dot{\eta}) dt.$$

Integration by parts moves the time derivative from the arbitrary variation to $L_{\dot{x}}$, producing a boundary term and

$$\delta \mathcal{S} = \int \left[L_x - \frac{d}{dt} L_{\dot{x}} \right] \eta dt.$$

Because $\eta(t)$ is arbitrary in the interior,

$$\boxed{\frac{d}{dt} L_{\dot{x}} - L_x = 0.}$$

For many coordinates,

$$\boxed{\frac{d}{dt} \frac{\partial L}{\partial \dot{q}_i} - \frac{\partial L}{\partial q_i} = 0.}$$

The conceptual chain is

$$\boxed{\text{variation of a path} \longrightarrow \text{chain rule} \longrightarrow \text{integration by parts} \longrightarrow \text{boundary term} + \text{local equation}.}$$

In the next lecture we will exploit the fact that the coordinates q_i need not be Cartesian. We will use generalized coordinates to describe constrained motion and see how the Lagrangian formulation can eliminate constraint forces from the problem before we begin solving the equations of motion.