Variational Principles

Shing Tak Lam

May 11, 2022

Contents

1 Functionals and multivariate calculus

1.1 Function spaces and functionals

In this course, we will be trying to find the optimum value of a functional[1]. But to be able to do this, we will need to first introduce some spaces[2] of functions.

Definition 1.1 (C^k functions)
Let A be a subset of R^n. Then for $k \in [0, \infty]$, define the space of k-times continuously differentiable functions.

$$C^k(A) := \{f : A \to \mathbb{R} : f \ k\text{-times continuously differentiable}\}$$

We write $C(A) = C^0(A)$ for continuous functions, and

$$C^k_{(\alpha,\beta)}(A) := \{f : A \to \mathbb{R} : f \ k\text{-times continuously differentiable}, f(\alpha) = f(\beta)\}$$

[1] To be introduced later
[2] These are *vector* spaces, as well as *normed* spaces with an appropriate domain/metric/norm.

for the space of functions with fixed end points.

Definition 1.2 (Functional)
A functional is a (real-valued) function from a space of functions.

1.2 Multivariate calculus

In this section, suppose $f \in C^2(\mathbb{R}^n)$. That is, f is twice differentiable with continuous second derivatives.

Definition 1.3 (Stationary point)
$a \in \mathbb{R}^n$ is a stationary point of f if

$$\nabla f(a) = (\partial_1 f, \cdots, \partial_n f)(a) = 0$$

where $\partial_i f = \partial f / \partial x_i$.

Definition 1.4 (Hessian)
The Hessian of f is the matrix $H_{ij} = \partial_i \partial_j f$.

Theorem 1.5 (Multivariate Taylor Theorem).

$$f(a+h) = f(a) + h \cdot \nabla f(a) + \frac{1}{2} h^\top H h + \mathcal{O}\left(\|h\|^2\right)$$

Proposition 1.6 (Classification of stationary points). Suppose (without loss of generality), $\nabla f(0) = 0$. Let λ_i be the eigenvalues of H. Then

(i) If all $\lambda_i > 0$, then f has a local minimum at 0.

(ii) If all $\lambda_i < 0$, then f has a local maximum at 0.

(iii) If some $\lambda_i > 0$, and some $\lambda_i < 0$, then f has a saddle point at 0.

(iv) If any $\lambda_i = 0$, then we need higher order terms.

Proof. Since f has continuous second order derivatives, H is a symmetric matrix. Thus it can be diagonalised by an orthogonal transformation. Hence without loss of generality, assume H diagonal. Then the multivariate Taylor theorem gives us that

$$f(x) - f(0) = \underbrace{x \cdot \nabla f(0)}_{=0} + \frac{1}{2} x^t H x + \mathcal{O}\left(||x||^2\right) = \sum \lambda_i x_i^2 + \mathcal{O}\left(||x||^2\right)$$

The result follows. □

1.3 Lagrange multipliers

Sometimes, we have constraints which limit domain which we are optimising over. We can use the method of Lagrange multipliers for this.

Definition 1.7 (Lagrange multipliers)
Suppose we have $f : \mathbb{R}^n \to \mathbb{R}$, and functions $(g_\alpha)_{\alpha=1}^k : \mathbb{R}^n \to \mathbb{R}$, and we wish to optimise f subject to the constraints $g_\alpha = 0$ for all α. Define $h : \mathbb{R}^n \times \mathbb{R}^k \to \mathbb{R}$ by

$$h(x, \lambda) = f(x) - \sum_{\alpha=1}^{k} \lambda_\alpha g_\alpha(x)$$

Then the optimum value can be found by considering $\nabla h = 0$. λ is known as the Lagrange multiplier.

2 Convexity

Definition 2.1 (Convex set)
A subset $S \subseteq \mathbb{R}^n$ is convex if for all $x, y \in S$, $t \in [0, 1]$, we have that

$$(1 - t) \cdot x + t \cdot y \in S$$

Definition 2.2 (Convex function)
For a convex set S, function $f : S \to \mathbb{R}$ is convex if for all $t \in [0, 1]$,

$$f((1 - t) \cdot x + t \cdot y) \leq (1 - t) \cdot f(x) + t \cdot f(y)$$

We say that f is strictly convex if the inequality is strict on $(0, 1)$.

Remark 2.3. We can define f is (strictly) concave if $-f$ is (strictly) convex.

2.1 Conditions for convexity

Proposition 2.4. Suppose f is differentiable. Then f is convex if and only if for all x, y,

$$f(y) \geq f(x) + (y - x) \cdot \nabla f(x) \tag{*}$$

Proof. Suppose (*) holds. Let $z_t = (1 - t) \cdot x + t \cdot y$. Then we have that

$$f(x) \geq f(z_t) + (x - z_t)\nabla f(z_t) \quad \text{and} \quad f(y) \geq f(z_t) + (y - z_t)\nabla f(z_t)$$

Then we get that

$$(1 - t) \cdot f(x) + t \cdot f(y) \geq f(z_t)$$

Which means that f is convex. Suppose instead that f is differentiable and convex. Let

$$h(t) = (1 - t) \cdot f(x) + t \cdot f(y) - f((1 - t) \cdot x + t \cdot y)$$

Then h is differentiable, with $h'(0) = -f(x) + f(y) - (y - x) \cdot \nabla f(x)$. So suffices to show that $h'(0) \geq 0$. $h(0) = 0$, and for $t \in (0, 1)$, we have that

$$\frac{h(t) - h(0)}{t} = \frac{h(t)}{t} \geq 0$$

Taking the limit $t \to 0$, we get the required result. □

Corollary 2.5. If f is convex and differentiable, with $\nabla f(x) = 0$. Then f has a global minimum at x.

Proof. If $\nabla f(x) = 0$, then x is a local minimum. By convexity, a local minimum is a global minimum. □

Proposition 2.6. Suppose f differentiable, with

$$(\nabla f(y) - \nabla f(x)) \cdot (y - x) \geq 0 \tag{**}$$

for all x, y. Then f is convex.

Proof. We have that

$$\begin{aligned} f(y) - f(x) &= [f((1-t)\cdot x + t \cdot y)]_0^1 \\ &= \int_0^1 \frac{\mathrm{d}}{\mathrm{d}t}(f((1-t)\cdot x + t\cdot y))\mathrm{d}t \\ &= \int_0^1 (y-x)\cdot \nabla f((1-t)\cdot x + t \cdot y)\mathrm{d}t \end{aligned}$$

Which means that

$$f(y) - f(x) - (y-x)\cdot \nabla f(x) = \int_0^1 (y-x)\cdot(\nabla f((1-t)\cdot x + t\cdot y) - \nabla f(x))\mathrm{d}t$$

Setting $y = (1-t)x + ty$ in (**), we get that $t(y-x)\cdot(\nabla f((1-t)\cdot x + t\cdot y) - \nabla f(x)) \geq 0$, which means that

$$f(y) \geq f(x) + (y-x)\cdot \nabla f(x)$$

So f is convex. □

Remark 2.7. The converse is also true, but not proven here.

Proposition 2.8. Suppose f twice differentiable, then f is convex if and only if all eigenvalues of the Hessian are nonnegative. Furthermore, it is strictly convex if and only if all are positive.

Proof. Suppose f convex. Then we have that for any h,

$$h \cdot (\nabla f(x+h) - \nabla f(x)) \geq 0$$

Furthermore, from the multivariate Taylor expansion, we have that

$$\frac{\partial f}{\partial x_i} = \frac{\partial f}{\partial x_i} + h_j \frac{\partial^2 f}{\partial x_i \partial x_j}(x) + \mathcal{O}\left(\|h\|^2\right) = \frac{\partial f}{\partial x_i} + h_j H_{ij}(x) + \mathcal{O}\left(\|h^2\|\right)$$

Hence taking the scalar product with h, we find that

$$h_i H_{ij} h_j + \mathcal{O}\left(\|h\|^3\right) \geq 0$$

Hence all eigenvalues are nonnegative. □

2.2 Legendre transform

Definition 2.9 (Legendre transform)
Let $f : \mathbb{R}^n \to \mathbb{R}$, the Legendre transform of f is

$$f^*(p) = \sup_x (p \cdot x - f(x))$$

where the domain of f^* is the set of all p such that the supremum exists.

Proposition 2.10. The domain of f^* is a convex set, and f^* is a convex function.

Proof.

$$\begin{aligned} f^*((1-t) \cdot p + t \cdot q) &= \sup_x ((1-t) \cdot p \cdot x + t \cdot q \cdot x - f(x)) \\ &= \sup_x ((1-t) \cdot (p \cdot x - f(x)) + t \cdot (q \cdot x - f(x))) \\ &\leq (1-t) \cdot \sup_x (p \cdot x - f(x)) + t \cdot \sup_x (q \cdot x - f(x)) \\ &= (1-t) \cdot f^*(p) + t \cdot f^*(q) \end{aligned}$$

Which both shows that the domain is convex, as well as f^* being convex. □

Since the properties of the Legendre transform on non-convex functions are not as well behaved, we now look at some properties assuming the function f is convex. First, we need the following lemma.

Lemma 2.11. Let $f : S \to \mathbb{R}$ be convex, and $p \in \mathbb{R}^n$ fixed. Then the function

$$g(x) = f(x) - x \cdot p$$

is also convex.

Thus, for a convex function, the Legendre transform is simply the maximum of a concave function, which must occur either on ∂S (if $\nabla g \neq 0$ on S), or at a point where the gradient is zero. Thus, we get that

Proposition 2.12. For a convex differentiable function $f : S \to \mathbb{R}$, we have that

$$p = \nabla f$$

Proof. We have that $\nabla g(x) = 0 \iff p = \nabla f(x)$. □

2.3 Legendre transform in Thermodynamics

Definition 2.13 (Quantities in Thermodynamics)
Define the following quantities

- Pressure - P
- Volume - V
- Temperature - T
- Entropy - S

Definition 2.14 (Energies in Thermodynamics)
Define the following energies

- Internal energy - $U(S, V)$
- Helmholtz free energy - $F(T, V) = -U^*(T, V)$
- Enthalpy - $H(S, P) = -U^*(S, P)$

3 Euler-Lagrange equations

We wish to extremise the functional

$$F[y] = \int_\alpha^\beta f(x, y, y')\mathrm{d}x$$

subject to regularity conditions, as well as the boundary conditions $y(\alpha) = y_1$ and $y(\beta) = y_2$. To do this, we will first need a lemma.

Lemma 3.1 (Fundamental lemma of calculus of variations). Ler $g : [\alpha, \beta] \to \mathbb{R}$ be continuous, and suppose for all $\eta : [\alpha, \beta] \to \mathbb{R}$ continuous, $\eta(\alpha) = \eta(\beta) = 0$, we have that

$$\int_\alpha^\beta g(x)\eta(x)\mathrm{d}x = 0$$

Then $g \equiv 0$ on $[\alpha, \beta]$.

Proof. We will prove the contrapositive. Suppose $g \not\equiv 0$. Say (without loss of generality) $g(y) > 0$ for some $y \in [\alpha, \beta]$. Then by continuity, we have $[a, b]$ such that $g(x) > g(y)/2 > 0$ for all $x \in [a, b]$. Let η be any continuous function which is zero outside $[a, b]$ and positive in (a, b). Then

$$\int_\alpha^\beta g(x)\eta(x)\mathrm{d}x = \int_a^b g(x)\eta(x)\mathrm{d}x > 0$$

□

Theorem 3.2 (Euler-Lagrange equation). A necessary condition for y to be an extremum of the functional

$$F[y] = \int_\alpha^\beta f(x, y, y')\mathrm{d}x$$

is for it to satisfy the Euler-Lagrange equation

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

Proof. Fix y, and we consider a small perturbation $y + \varepsilon\eta$, where $\varepsilon \in \mathbb{R}$ small and η satisfying the regularity conditions, as well as the boundary condition $\eta(\alpha) = \eta(\beta) = 0$. Substituting, and expanding as a function of ε, we find that

$$\begin{aligned} F[y + \varepsilon\eta] &= \int_\alpha^\beta f(x, y + \varepsilon\eta, y' + \varepsilon\eta')\mathrm{d}x \\ &= F[y] + \varepsilon \int_\alpha^\beta \left(\frac{\partial f}{\partial y}\eta + \frac{\partial f}{\partial y'}\eta'\right) \mathrm{d}x + \mathcal{O}(\varepsilon^2) \end{aligned}$$

For this to be an extremum, we want $\mathrm{d}F/\mathrm{d}\varepsilon = 0$. Integrating by parts, we find that

$$
\begin{aligned}
0 &= \frac{\mathrm{d}F}{\mathrm{d}\varepsilon} \\
&= \int_\alpha^\beta \left(\frac{\partial f}{\partial y}\eta + \frac{\partial f}{\partial y'}\eta' \right) \mathrm{d}x \\
&= \int_\alpha^\beta \left(\frac{\partial f}{\partial y}\eta - \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial f}{\partial y'} \right)\eta \right) \mathrm{d}x + \underbrace{\left[\frac{\partial f}{\partial y'}\eta \right]_\alpha^\beta}_{=0} \\
&= \int_\alpha^\beta \left(\frac{\partial f}{\partial y} - \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial f}{\partial y} \right) \right) \eta \mathrm{d}x
\end{aligned}
$$

Since this must hold for all η, applying the fundamental lemma we get the required result. □

Remark 3.3. Euler-Lagrange is a second order ODE in y with boundary conditions.

3.1 First Integrals

In some special cases, E-L reduces to a first order ODE.

Corollary 3.4. If in addition, $\partial f / \partial y = 0$, then a necessary condition is

$$
\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial f}{\partial y'} \right) = 0 \iff \frac{\partial f}{\partial y'} = \mathrm{const}
$$

Corollary 3.5. If instead $\partial f / \partial x = 0$, then a necessary condition is

$$
f - y'\frac{\partial f}{\partial y'} = \mathrm{const}
$$

Proof.

$$
\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}x}\left(f - y'\frac{\partial f}{\partial y'} \right) &= \frac{\partial f}{\partial x} + \frac{\partial f}{\partial y}y' + \frac{\partial f}{\partial y'}y'' - y''\frac{\partial f}{\partial y'} - y'\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial f}{\partial y'} \right) \\
&= \frac{\partial f}{\partial x} + y'\left(\frac{\partial f}{\partial y} - \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial f}{\partial y'} \right) \right) \\
&= 0
\end{aligned}
$$

□

3.2 Constrained E-L Equations

Suppose we wished to extremise

$$
F[y] = \int_\alpha^\beta f(x, y, y')\mathrm{d}x
$$

subject to the constraint that

$$
G[y] = \int_\alpha^\beta g(x, y, y')\mathrm{d}x = 0
$$

Then we can use Lagrange multipliers, and extremise

$$
\Phi[y, \lambda] = F[y] - \lambda G[y]
$$

This then gives us

Proposition 3.6. A necessary condition for y to be an extremum of the functional F, subject to the condition $G[y] = 0$ as above, is

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial}{\partial y'}(f - \lambda g)\right) - \frac{\partial}{\partial y}(f - \lambda g) = 0$$

Proof. Euler-Lagrange with the functional Φ. □

3.2.1 Sturm-Liouville

One important example of this is known as the Sturm-Liouville operator.

Example 3.7 (Sturm-Liouville)
Let $\rho(x) > 0$ on a set $[\alpha, \beta]$, and $\sigma(x)$ is any function. Then we wish to extremise the functional

$$F[y] = \int_\alpha^\beta \rho(y')^2 + \sigma y^2 \mathrm{d}x$$

subject to the boundary conditions on y, as well as

$$G[y] = \int_\alpha^\beta y^2 \mathrm{d}x = 1$$

Define the functional $\Phi[y, \lambda] = F[y] - \lambda(G[y] - 1)$. Then the Euler-Lagrange equations give us

$$\frac{\partial f}{\partial y'} = 2\rho y' \quad \text{and} \quad \frac{\partial h}{\partial y} = 2\sigma y - 2\lambda y$$

Which means that

$$\mathcal{L}[y] = -\frac{\mathrm{d}}{\mathrm{d}x}\left(\rho y'\right) + \sigma y = \lambda y$$

where $\mathcal{L}$ is known as the Sturm-Liouville operator. Since $\mathcal{L}$ is linear, we see that this is an eigenvalue problem.

3.3 Several dependent variables

Theorem 3.8 (Euler-Lagrange with several dependent variables)**.** Suppose we have $y : \mathbb{R} \to \mathbb{R}^n$, a necessary condition for y to be an extremum of the functional

$$F[y] = \int_\alpha^\beta f(x, y_1, \dots, y_n, y_1', \dots, y_n')\mathrm{d}x$$

is

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial f}{\partial y_i'}\right) - \frac{\partial f}{\partial y_i} \quad \text{for} \quad i = 1, \dots, n$$

Proof. Apply perturbation $\eta : [\alpha, \beta] \to \mathbb{R}^n$ as in the proof of the single variable case. □

Like with the single variable case, we have first integrals

Corollary 3.9. For any $j \in \{1, \dots, n\}$, if $\partial f / \partial y_j = 0$, then

$$\frac{\partial f}{\partial y_j'} = \text{const}$$

Corollary 3.10. If $\partial f / \partial x = 0$, then

$$f - \sum_i y_i' \frac{\partial f}{\partial y_i'} = \text{const}$$

3.4 Several independent variables

Now, we consider the most general case, where we have $\phi : \mathbb{R}^n \to \mathbb{R}^m$. Instead of an interval in $\mathbb{R}$, consider a subset $\mathcal{D} \subseteq \mathbb{R}^n$, and the boundary conditions will by imposed on the boundary $\partial\mathcal{D}$.

Theorem 3.11. A necessary condition for $\phi : \mathbb{R}^n \to \mathbb{R}^m$ to be an extremum for the functional

$$F[\phi] = \int_{\mathcal{D}} f(x, \phi, \phi')\mathrm{d}V$$

where ϕ' is the (Fréchet) derivative of ϕ, i.e. a matrix of partial derivatives $\partial\phi_i / \partial x_j$, is if

$$\frac{\partial f}{\partial \phi_i} - \sum_{j=1}^{n} \frac{\partial}{\partial x_j}\left(\frac{\partial f}{\partial(\partial_j \phi_i)}\right) = 0 \quad \text{for} \quad i = 1, \ldots, m$$

Proof. Apply perturbation $\phi + \varepsilon\eta$ as before, where we have the boundary condition that $\eta = 0$ on $\partial\mathcal{D}$, we have that

$$F[\phi + \varepsilon\eta] - F[\phi] = \varepsilon \sum_{i=1}^{m} \int_{\mathcal{D}} \left(\frac{\partial f}{\partial \phi_i}\eta_i + \sum_{j=1}^{n} \frac{\partial f}{\partial(\partial_j \phi_i)}\partial_j \eta_i\right) \mathrm{d}V$$

We must have that each of the integrals that we are summing over are zero, as we can perturb (further) only in one of the components. Hence we may assume without loss of generality that $m = 1$, and we will drop the i subscripts. Then

$$\begin{aligned}
F[\phi + \varepsilon\eta] - F[\phi] &= \varepsilon \int_{\mathcal{D}} \eta\frac{\partial f}{\partial \phi} + \sum_j \frac{\partial f}{\partial(\partial_j \phi)}\partial_j \eta \mathrm{d}V \\
&= \varepsilon \int_{\mathcal{D}} \eta\frac{\partial f}{\partial \phi} + \underbrace{\nabla \cdot \left(\eta \cdot \left(\frac{\partial f}{\partial \partial_1 \phi}, \ldots, \frac{\partial f}{\partial \partial_n \phi}\right)\right)}_{=0 \text{ by Divergence Theorem as } \eta = 0 \text{ on } \partial\mathcal{D}} - \eta\nabla \cdot \left(\frac{\partial f}{\partial \partial_1 \phi}, \ldots, \frac{\partial f}{\partial \partial_n \phi}\right)\mathrm{d}V \\
&= \varepsilon \int_{\mathcal{D}} \left(\frac{\partial f}{\partial \phi} - \sum_{j=1}^{n} \frac{\partial f}{\partial(\partial_j \phi)}\right)\eta \mathrm{d}V
\end{aligned}$$

Applying the fundamental lemma gives the required result. □

3.5 Higher order derivatives

Theorem 3.12. A necessary condition for y to be an extremum of the functional

$$F[y] = \int_{\alpha}^{\beta} f(x, y, y', \ldots, y^{(n)})\mathrm{d}x$$

is if

$$\sum_{i=0}^{n}(-1)^i \frac{\mathrm{d}^n}{\mathrm{d}x^n}\left(\frac{\partial f}{\partial y^{(n)}}\right) = 0$$

Proof. Apply perturbation $y + \varepsilon\eta$, where we require $\eta = \eta' = \cdots = \eta^{(n-1)} = 0$ at α, β. Then expanding we find that

$$F[y+\varepsilon\eta] - F[y] = \varepsilon \int_\alpha^\beta \left(\frac{\partial f}{\partial y}\eta + \frac{\partial f}{\partial y}\eta' + \cdots + \frac{\partial f}{\partial y^{(n)}}\eta^{(n)}\right) \mathrm{d}x + \mathcal{O}\left(\varepsilon^2\right)$$

Leave the first term as is, integrate the second one by parts, and so on, where we integrate the last one by parts n times, we get the required result. □

4 Dynamics

4.1 Least action principle

Definition 4.1 (Lagrangian)
For a particle in $\mathbb{R}^3$ with kinetic energy T and potential energy V, the Lagrange is

$$L(x, \dot{x}, t) = T - V$$

Definition 4.2 (Action)
The action of a particle is

$$S[x] = \int_{t_1}^{t_2} L(x, \dot{x}, t)\mathrm{d}t$$

Proposition 4.3 (Least action principle). The motion of a particle is such that $S[x]$ is stationary. That is, x satisfies the Euler-Lagrange equations.

4.2 Noether's theorem

Definition 4.4 (Continuous symmetry of Lagrangian)
Suppose we have a functional

$$F[y] = \int_\alpha^\beta f(x, y_1, \ldots, y_n, y_1', \ldots, y_n')\mathrm{d}x$$

Then a one parameter transformation $Y(x, s)$ of y, where $Y(x, 0) = y(x)$ is called a continuous symmetry of f if

$$\frac{\mathrm{d}}{\mathrm{d}s} f(x, Y_1(x,s), \ldots, Y_n(x,s), Y_1'(x,s), \ldots, Y_n'(x,s))$$

where primes denote derivatives with respect to x.

Theorem 4.5 (Noether's theorem). Given a continuous symmetry $Y(x, s)$ of f,

$$\sum_i \frac{\partial f}{\partial y_i}\frac{\partial Y_i}{\partial s}\bigg|_{s=0}$$

is a first integral of the Euler-Lagrange equations. That is, it is a conserved quantity.

Proof. In this proof we will use the summation convention.

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial f}{\partial y_i}\frac{\partial Y_i}{\partial s}\right)\Bigg|_{s=0} &= \left(\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial f}{\partial y_i'}\right)\frac{\partial Y_i}{\partial s} + \frac{\partial f}{\partial y_i'}\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial Y_i}{\partial s}\right)\right)\Bigg|_{s=0} \\
&= \frac{\partial f}{\partial y_i}\frac{\partial Y_i}{\partial s}\Bigg|_{s=0} + \frac{\partial f}{\partial y_i'}\frac{\partial Y_i'}{\partial s}\Bigg|_{s=0} \\
&= \frac{\mathrm{d}f}{\mathrm{d}s}\Bigg|_{s=0} \\
&= 0
\end{aligned}$$

□

4.3 Hamilton's equations

Definition 4.6 (Hamiltonian)
Let $L = L(q, \dot{q}, t) = T - V$ be the Lagrangian of a system. Define the Hamiltonian

$$H(q, p, t) = \sup_v (p \cdot v - L(q, v, t))$$

to be the Legendre transform of L with respect to $v = \dot{q}$. p is called the generalised momentum.

Theorem 4.7 (Hamilton's equations)**.** Suppose that L is convex and differentiable. Then we have that

$$\dot{q}_i = \frac{\partial H}{\partial p_i} \quad \text{and} \quad \dot{p}_i = -\frac{\partial H}{\partial q_i} \quad \text{and} \quad \frac{\partial H}{\partial t} = -\frac{\partial L}{\partial t}$$

Proof. First we note that (using the summation convention)

$$H = p_i v_i - L(q_i, v_i, t)$$

where $v_i(p)$ is a solution to $p_i = \partial L / \partial \dot{q}_i$. So we have that

$$\begin{aligned}
\frac{\partial H}{\partial q_i}\mathrm{d}q_i + \frac{\partial H}{\partial p_i}\mathrm{d}p_i + \frac{\partial H}{\partial t}\mathrm{d}t = \mathrm{d}H& \\
&= p_i \mathrm{d}v_i + v_i \mathrm{d}p_i - \frac{\partial L}{\partial q_i}\mathrm{d}q_i - \frac{\partial L}{\partial \dot{q}_i}\mathrm{d}v_i - \frac{\partial L}{\partial t}\mathrm{d}t \\
&= v_i \mathrm{d}p_i - \frac{\partial L}{\partial q_i}\mathrm{d}q_i - \frac{\partial L}{\partial t}\mathrm{d}t
\end{aligned}$$

Writing $\dot{q}_i = v_i$ and equating differentials gives the required result. □

Alternative proof. Consider the functional

$$S[q, p] = \int_{t_1}^{t_2} \dot{q}_i p_i - H(q, p, t)\mathrm{d}t$$

Extremising and considering the Euler-Lagrange equations gives the required result. □

5 Second variation

Definition 5.1 (Second variation)
For a functional

$$F[y] = \int_{\alpha}^{\beta} f(x, y, y')\mathrm{d}x$$

Define the second variation

$$\delta^2 F[y, \eta] = \frac{1}{2}\int_{\alpha}^{\beta} \eta^2 \frac{\partial^2 f}{\partial y^2} + (\eta')2\frac{\partial^2 f}{\partial (y')^2} + \frac{\mathrm{d}}{\mathrm{d}x}\left(\eta^2\right)\frac{\partial^2 f}{\partial y \partial y'}\mathrm{d}x$$

Proposition 5.2. Suppose y is a solution to the Euler-Lagrange equation. Then expanding $F[y + \varepsilon\eta]$ to $\mathcal{O}\left(\varepsilon^3\right)$, we get that

$$F[y + \varepsilon\eta] - F[y] = \varepsilon^2\, \delta^2 F[y, \eta] + \mathcal{O}\left(\varepsilon^3\right)$$

Proposition 5.3. Let

$$P = \frac{\partial^2 f}{\partial (y')^2} \quad \text{and} \quad Q = \frac{\partial^2 f}{\partial y^2} - \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial^2 f}{\partial y \partial y'}\right)$$

Then

$$\delta^2 F[y, \eta] = \frac{1}{2}\int_{\alpha}^{\beta} Q\eta^2 + P(\eta')^2\mathrm{d}x$$

Proof. Integrate $\frac{\mathrm{d}}{\mathrm{d}x}\left(\eta^2\right)\frac{\partial^2 f}{\partial y \partial y'}$ by parts, and use the boundary condition that $\eta = 0$ at α, β. □

Proposition 5.4 (Legendre condition)**.** If $y(x)$ is a solution to the Euler-Lagrange equation, and $Q\eta^2 + P(\eta')^2 > 0$ for all η such that $\eta = 0$ at α, β, then y is a local minimum of $F[y]$.

Proposition 5.5.

$$\delta^2 F[y] = \frac{1}{2}\int_{\alpha}^{\beta} \eta\mathcal{L}[\eta]\mathrm{d}x$$

where

$$\mathcal{L}[\eta] = -(P\eta')' + Q\eta$$

is the Sturm-Liouville operator.

Proof. Integration by parts, and use boundary conditions. □

Corollary 5.6. If there exists $\omega \in \mathbb{R}$, η such that

$$\mathcal{L}[\eta] = -\omega^2\eta$$

Then y cannot be a minimiser.

Proof. Substituting we get that

$$\delta^2 F[y] = -\frac{1}{2}\omega^2 \int_\alpha^\beta \eta^2 \mathrm{d}x < 0$$

□

5.1 Jacobi condition

Theorem 5.7 (Jacobi accessory equation)**.** Suppose $P|_y > 0$ and we have u nonzero such that

$$-(Pu')' + Qu = 0$$

Then we must have that $\delta^2 F[y, u] > 0$, and y is a local minimum.

Proof. Let ϕ be any function. Note that we have

$$0 = \int_\alpha^\beta (\phi\eta^2)' \mathrm{d}x = \int_\alpha^\beta \phi'\eta^2 + 2\eta\eta'\phi \mathrm{d}x$$

Adding this to the second variation we find that

$$\delta^2 F[y] = \frac{1}{2}\int_\alpha^\beta P(\eta')^2 + 2\eta\eta'\phi + (Q + \phi')\eta \mathrm{d}x$$

Completing the square, we have that

$$\delta^2 F[y] = \frac{1}{2}\int_\alpha^\beta P\left(\eta' + \frac{\phi}{P}\eta\right)^2 + \left(Q + \phi' - \frac{\phi^2}{P}\right)\eta^2 \mathrm{d}x$$

We can make this postive if we can solve

$$Q + \phi' - \frac{\phi^2}{P} = 0 \tag{*}$$

Since for

$$\int_\alpha^\beta P\left(\eta' + \frac{\phi}{P}\eta\right)^2 \mathrm{d}x = 0$$

We must have that $\eta \equiv 0$. Now to solve $*$, we can transform it into a second order linear equation by substituting $\phi = -Pu'/u$, and we get the required result. □