CLASSIC THEORY OF CALCULUS OF VARIATION

LONG CHEN

ABSTRACT. We present the first-order condition: the Euler–Lagrange equation, and various second-order conditions: the Legendre condition, the Jacobi condition, and the Weierstrass condition. We also discuss weak and strong minima. For the classical theory, the presentation is limited to one-dimensional functions.

CONTENTS

1. PROBLEM FORMULATION

The classical calculus of variation problem is

$$\inf_{x\in\mathcal{M}} \int_a^b L(t, x(t), x'(t))\ \mathrm{d}t, \tag{1}$$

with

- $\mathcal{M} = \{x \in C^1(a,b) \mid x(a) = A,\ x(b) = B\}$;
- $L(t,x,v) : \mathbb{R}^3 \to \mathbb{R}$ is sufficiently smooth, e.g., of class C^2.

In the notation $L(t,x,v)$, the variables (t,x,v) are treated as three independent variables, often referred to as time, state, and velocity. However, in $L(t,x(t),x'(t))$, these variables are coupled through the function $x(t)$.

A function $x_* \in \mathcal{M}$ is called a global minimizer if

$$I(x_*) \le I(x) \quad \forall x \in \mathcal{M},$$

Date: June 4, 2025.

and a local minimizer if

$$I(x_*) \leq I(x) \quad \forall x \in \mathcal{N}(x_*),$$

where the neighborhood $\mathcal{N}(x_*)$ will be defined more precisely in Definition 4.1. If the inequality $\leq$ is replaced by a strict inequality $<$, then x_* is called a strict minimizer.

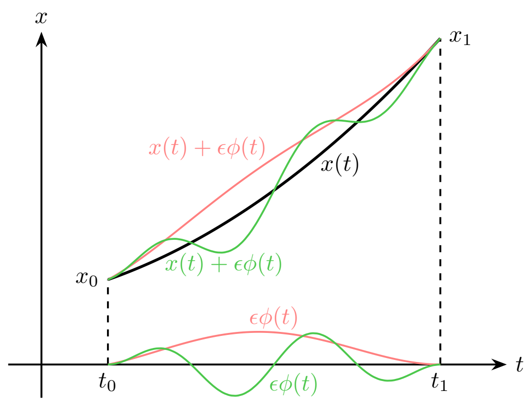


FIGURE 1. Variations of a curve.

We introduce a variation of a function. Let ϕ be a test function in $\mathcal{M}_0$, satisfying certain boundary conditions such that if $x \in \mathcal{M}$, then $x + \epsilon\phi \in \mathcal{M}$ for small ϵ. For the example considered,

$$\mathcal{M}_0 = C_0^1(a, b) := \{\phi \in C^1(a, b) \mid \phi(a) = \phi(b) = 0\}.$$

The term $\epsilon\phi$ is called a variation of x, and the scalar ϵ indicates that the perturbation is small.

Define the function

$$f(\epsilon) := I(x_* + \epsilon\phi).$$

We are now reduced to a standard problem in calculus: a curve x_* is an interior local minimizer of $I(\cdot)$ if and only if 0 is a local minimizer of the function $f(\epsilon)$. Thus, the first-order optimality condition for an extremal curve x_* of $I(x)$ is

$$f'(0) = \frac{\mathrm{d}}{\mathrm{d}\epsilon} I(x + \epsilon\phi)\Big|_{\epsilon=0} = 0. \tag{2}$$

Moreover, we have the following sufficient and necessary conditions for 0 to be a local and interior minimizer of $f(\epsilon)$:

$$f'(0) = 0, \; f''(0) > 0$$
$$\Downarrow$$
$$x_* \text{ is an interior local minimizer}$$
$$\Downarrow$$
$$f'(0) = 0, \; f''(0) \geq 0$$

Typically, we are concerned with interior minimizers. When $\mathcal{M}$ is bounded, the minimizer may lie on the boundary. In such cases, the first-order optimality condition becomes a variational inequality, which will be discussed in a separate lecture note.

2. THE EULER-LAGRANGE EQUATION

By the chain rule, we obtain the variational form of Euler-Lagrange equation, where $(\cdot,\cdot)$ is the stand L^2-inner product of function space $L^2(a,b)$.

2.1. **Weak and Strong form of E-L Equation.** By the chain rule, we obtain from (2) the weak form of E-L equation.

Variational (weak) form of Euler-Lagrange equation

$$(L_v(t,x,x'),\phi') + (L_x(t,x,x'),\phi) = 0 \quad \forall \phi \in H_0^1(a,b). \tag{3}$$

In the context of the first-order variation $\delta I(x,\phi)$, it is sufficient for x and ϕ to belong to the class C^1. However, the C^1 condition is relatively strong, since it demands that derivatives exist in the classical sense as pointwise limits of difference quotients.

Example 2.1. Consider the following example (Example 15.1 p307 in [2]):

$$\min_{x\in\mathcal{M}} I(x) := \int_{-1}^{1} x^2(x'-1)^2\,\mathrm{d}t, \qquad \text{with } x(-1)=0, x(1)=1.$$

Obviously $I(x)\geq 0$ and $x_* = \max\{t,0\}$ (ReLU function) is a global minimum which is not in C^1. One can easily construct a sequence of C^1 curves x_n s.t. $I(x_n)\to 0$ and $x_n \to x_*$ pointwisely. That is we find a minimizing sequence $\{x_k\}\subset C^1$ but the limit is out of C^1. In other words, C^1 is not complete under the norm $\|\cdot\|_{1,\infty}$.

In many applications, it is natural to consider a broader class of functions—those that are continuous and piecewise smooth. These functions may have corners or kinks and thus fail to be classically differentiable everywhere. This motivates the introduction of a more flexible notion of differentiability, such as the concept of *weak derivatives*.

The space for the test function ϕ can be relaxed to $H_0^1(a,b) := \overline{C_0^1(a,b)}^{\|\cdot\|_1}$, where

$$(u,v)_1 := (u,v) + (u',v') = \int_a^b uv + u'v'\,\mathrm{d}t,$$

and $\|\cdot\|_1$ is the induced norm. In the variational form (3), the solution (if exists) $x \in H^1(a,b) = \overline{C^1(a,b)}^{\|\cdot\|_1}$. Functions in $H^1(a,b)$ possess weak derivatives, which are defined through action rather than the classical notion of derivative as a pointwise limit of the difference quotient. More precisely, for a function $x\in L^1_{\rm loc}(a,b)$, if there exists a function $p\in L^1_{\rm loc}(a,b)$ satisfying

$$(p,\phi) = -(x,\phi') \quad \text{for all } \phi\in C_0^\infty(a,b),$$

then we say p is the weak derivative of x and still denoted by $p = x'$. Obviously when x is C^1, the weak derivative coincides with the classic derivative. For piecewise C^1 function and globally continuous functions, the weak derivative exists and equals to the piecewise derivative. It may not be differentiable in the classical sense, as demonstrated by examples such as the absolute value function and the ReLU function. In particular the piecewise linear and globally continuous function, which is the linear finite element space, will be useful to construct special test functions in the proof.

In general, for $1\leq p\leq\infty$, we introduce the Sobolev space $W^{1,p}(a,b) = \overline{C^1(a,b)}^{\|\cdot\|_{1,p}}$ and $W_0^{1,p}(a,b) = \overline{C_0^1(a,b)}^{\|\cdot\|_{1,p}}$, which are Banach spaces under the norm $\|x\|_{1,p} =$

$(\|x\|_{L^p}^p + \|x'\|_{L^p}^p)^{1/p}$ with standard modification for $p = \infty$. Since the functions are Lebesgue measurable, the norm $\|\cdot\|_\infty$ refers to the essential supremum.

By definition, for a continuous and piecewise smooth function, the weak derivative is equal to the piecewise derivative. It is important to note that continuity is still required to avoid discontinuities during integration by parts. In one dimension, if $x \in W^{1,1}(a,b)$, then we have the representation formula

$$x(t) = x(a) + \int_a^t x'(s)\,\mathrm{d}s. \tag{4}$$

The fact that x' exists almost everywhere implies that x is absolutely continuous. In particular, the space

$$W^{1,\infty}(a,b) \cong \mathrm{Lip}(a,b) = C^{0,1}(a,b).$$

Using integration by parts and noting that the boundary term disappears since ϕ vanishes on the boundary, we get the strong form of the E–L equation.

Strong form of Euler-Lagrange equation

$$-\frac{\mathrm{d}}{\mathrm{d}t}L_v(t,x(t),x'(t)) + L_x(t,x(t),x'(t)) = 0 \quad t \in [a,b] \quad \text{a.e.} \tag{5}$$

The test function ϕ is not present in (5) by using the fact: $\forall \phi \in L^2(a,b)$ holds $\rightarrow$ pointwise (a.e.) holds. If this were not the case, one could construct a test function ϕ near a point t that violates the weak form of the Euler-Lagrange equation (3), contradicting its validity for all $\phi \in H_0^1(a,b)$.

The strong form of the Euler–Lagrange equation holds for $x \in H^2(a,b)$, and (5) is a nonlinear second-order elliptic ODE. When the function is not smooth enough, we can write it in integral form.

Integral form of the Euler–Lagrange equation

$$-L_v(t,x(t),x'(t)) + \int_a^t L_x(s,x,x')\,\mathrm{d}s = \text{const}, \quad t \in [a,b] \quad \text{a.e.} \tag{6}$$

Proof. Start with the weak form (3) of the Euler–Lagrange equation and apply integration by parts to the lower-order term:

$$(L_x,\phi) = \left(\left(c + \int_a^t L_x\right)', \phi\right) = -\left(c + \int_a^t L_x, \phi'\right).$$

Then for any $c \in \mathbb{R}^n$ and any $\phi \in H_0^1(a,b)$, we have

$$\left(L_v - \int_a^t L_x - c, \phi'\right) = 0. \tag{7}$$

Choose a special test function

$$\phi(t) = \int_a^t \left[L_v - \int_a^s L_x - c\right]\,\mathrm{d}s,$$

which satisfies $\phi' = L_v - c - \int_a^t L_x$. Choose $c \in \mathbb{R}^n$ to satisfy the boundary conditions $\phi(a) = \phi(b) = 0$. Then (7) becomes $\|\phi'\| = 0$, and thus $\phi = 0$. □

Introduce $p(t) = L_v(t, x(t), x'(t))$, which is called the momentum or adjoint variable in physics, or the co-state in control theory. Then the integral form of the Euler–Lagrange equation can be rewritten as follows. This is the standard way to express the second-order ODE (5) as a first-order system.

Euler–Lagrange equation as a first-order system

$$p = L_v, \quad p' = L_x, \quad t \in [a, b] \quad \text{a.e.} \tag{8}$$

2.2. **Boundary and interface conditions.** When the admissible set is enlarged to $\mathcal{M} = H^1(a, b)$ without fixing the boundary condition, the test space also becomes $H^1(a, b)$ without the zero boundary condition. The variational form (3) still holds by first restricting the test function to the subspace $H_0^1(a, b)$. Then, applying integration by parts, we still obtain the strong form (5) of the E–L equation.

Next, choose $\phi \in H^1(a, b)$ and apply integration by parts again. Now, using the previously established E–L equation (5), we eliminate the volume contribution and conclude that the boundary term vanishes:

$$L_v(t, x, x')\, \phi\big|_a^b = 0, \ \forall \phi \in H^1(a, b) \quad \Longrightarrow \quad p(a) = p(b) = 0.$$

This shows that a Neumann-type boundary condition is built into the weak formulation of the E–L equation.

As an exercise, the reader is encouraged to verify the first Erdmann corner condition at an interface point c where the function is continuous but not differentiable.

First Erdmann Corner Condition

For an extremal $x \in C^0([a, b]) \cap C^1([a, c) \cup (c, b])$, i.e., $x(t)$ is continuous but may have a corner at $t = c \in (a, b)$. Then

$$p(t)|_{t=c^-} = p(t)|_{t=c^+}\,, \tag{9}$$

where $p(t) = L_v(t, x(t), x'(t))$ is the adjoint variable.

When L is independent of t, which is called *autonomous*, we have the conservation of Hamiltonian $H = px' - L$, which is known as the second Erdmann condition.

Second Erdmann condition for autonomous Lagrangian

$$x'(t)L_v(x(t), x'(t)) - L(x(t), x'(t)) = \text{const.} \tag{10}$$

In contrast to the strong form of the Euler-Lagrange equation given by (5), equation (10) is a first order nonlinear ODE and can be thought of as a first integral of (5). The validity of (10) can be easily verified by taking its derivative and using the strong form of the Euler-Lagrange equation.

2.3. **Regularity.** We relax $C^1[a,b]$ to $\mathrm{Lip}[a,b]$ to seek a minimum in a larger admissible function set. Under certain conditions, we can prove that the minimum found is indeed in the smaller space $C^1[a,b]$.

Theorem 2.2. *Let $x_* \in \mathrm{Lip}[a,b]$ satisfy the integral Euler–Lagrange equation (6), and assume that for almost every $t \in [a,b]$, the Lagrangian $L(t, x_*(t), v)$ is strictly convex in v. Then $x_* \in C^1[a,b]$.*

Proof. In the integral form

$$x(t) = x(a) + \int_a^t x'(s)\,\mathrm{d}s,$$

the weak derivative x' exists almost everywhere. If we can find a continuous representative of x', then the regularity is improved from $\mathrm{Lip}[a,b]$ to $C^1[a,b]$.

For a function $x_* \in \mathrm{Lip}[a,b]$, the derivative x_*' could be discontinuous at some point τ. Then $L_v(\tau, x_*(\tau), \ell_-)$ may not equal $L_v(\tau, x_*(\tau), \ell_+)$, where $\ell_\pm$ denote the left and right limits of x_*' at τ.

Using the integral form of the Euler–Lagrange equation (6), we have $L_v(\tau, x_*(\tau), \ell_-) = L_v(\tau, x_*(\tau), \ell_+)$ by taking two sequences converging to τ from the left and right. Since L is strictly convex in v, i.e., L_v is strictly increasing as a function of the third variable, it follows that $\ell_- = \ell_+$. Thus, x_*' is continuous at τ. □

In Example 2.1, the condition that 'L is strictly convex in v' is not valid, since $x_*(t) = 0$ holds on a set of nonzero measure. When L is C^2 and strictly convex, then $L_{vv} > 0$, and the Euler–Lagrange equation becomes an elliptic equation. In this case, a stronger regularity result $x \in C^2(a,b)$ can be established.

3. Second Order Conditions

In calculus, a critical point of the optimization problem $\min f(x)$ is a point where the first derivative vanishes, i.e., $f'(x) = 0$. Similarly, in the calculus of variations, a solution to the Euler–Lagrange equation is called an extremal (or extremal curve) of the functional. To determine whether an extremal is a local minimizer, second-order conditions must be considered. In particular, one must examine the behavior of the second variation of the functional in a neighborhood of the extremal. Throughout this section, we assume that the Lagrangian L is twice continuously differentiable, ensuring that the second variation exists and is well-defined.

3.1. **Second order variation.** We first consider the second order variation:

$$f''(0) = \delta^2 I(x,\phi) := \frac{\mathrm{d}^2}{\mathrm{d}\epsilon^2} I(x+\epsilon\phi)|_{\epsilon=0}.$$

In terms of the second order variation, we can write out conditions for 0 being a local minimizer of the single variable function $f(\epsilon)$.

- *Necessary conditions.* If $x_* \in \mathcal{M}$ is a local minimizer of $I(\cdot)$, then
  (1) $\delta I(x_*,\phi) = 0$, for all $\phi \in H_0^1(\Omega)$, and
  (2) $\delta^2 I(x_*,\phi) \geq 0$ for all $\phi \in H_0^1(\Omega)$.

- *Sufficient conditions.* If
  (1) $\delta I(x_*,\phi) = 0$, for all $\phi \in H_0^1(\Omega)$, and
  (2) there exists $\lambda > 0$ such that $\delta^2 I(x_*,\phi) \geq \lambda\|\phi\|_1^2$ for all $\phi \in H_0^1(\Omega)$,

then x_* is a strict local minimizer of $I(\cdot)$.

We then derive conditions without the test function ϕ.

3.2. **Hessian of the Lagrangian.** By direct calculation, we have

$$\delta^2 I(x,\phi) = (A\phi',\phi') + 2(B\phi',\phi) + (C\phi,\phi),$$

where

$$A = L_{vv}(t,x(t),x'(t)), \quad B = L_{xv}(t,x(t),x'(t)), \quad C = L_{xx}(t,x(t),x'(t)).$$

The Hessian matrix of $L(t,x,v)$ with respect to variable (v,x) at x_* is denoted by

$$\mathcal{H}^*(t) = \begin{pmatrix} A^* & B^* \\ B^* & C^* \end{pmatrix}. \tag{11}$$

Here * is used to emphasize it is evaluated at a particular extremal function x_* solving the E-L equation. Then the second order variation can be rewritten as

$$\delta^2 I(x_*,\phi) = \left(\mathcal{H}^* \begin{pmatrix} \phi' \\ \phi \end{pmatrix}, \begin{pmatrix} \phi' \\ \phi \end{pmatrix}\right).$$

Based on this formulation, we have the following sufficient condition.

A sufficient condition: Hessian of the Lagrangian is SPD

Suppose

(1) x_* satisfies the E-L equation of $I(\cdot)$;

(2) $\mathcal{H}^*(t) > 0$ for all $t \in [a,b]$;

then x_* is a strict local minimum.

Here, a symmetric $n \times n$ matrix $M > 0$ means $(Mv,v) \geq 0$ for all $v \in \mathbb{R}^n$, and $(Mv,v) = 0$ if and only if $v = 0$, or equivalently, $\lambda_{\min}(M) > 0$, where $\lambda_{\min}(M)$ is the minimum eigenvalue of M, and such M is called symmetric and positive definite (SPD).

Use the continuity of $\lambda_{\min}(\mathcal{H}^*(t))$ on the compact interval $[a,b]$, which follows from the assumption that L is twice continuously differentiable, we conclude that there exists a minimum value $\lambda_0 > 0$ s.t. $\min_{t\in[a,b]} \lambda_{\min}(\mathcal{H}^*(t)) > \lambda_0$. Then $\delta^2 I(x_*,\phi) \geq \lambda_0 \|\phi\|_1^2$.

The condition $\mathcal{H}^* \geq 0$ is, however, not necessary. Here is an example.

Example 3.1. *Consider $L(t,x,v) = v^2 - x^2$. Then $\mathcal{H}^* = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$. Consider zero boundary condition $x(a) = x(b) = 0$. Then $x_* = 0$ solves E-L equation. By the Poincaré inequality,*

$$\delta^2 I(x_*,\phi) = \|\phi'\|^2 - \|\phi\|^2 \geq \left(1 - \frac{(b-a)^2}{2}\right)\|\phi'\|^2, \quad \forall \phi \in H_0^1(a,b).$$

So for $(b-a)^2$ is smaller than 2, we conclude $x_* = 0$ is a local minimizer.

Exercise 3.2 (Poincaré inequality)**.** *For $\phi \in H_0^1(a,b)$, we have*

$$\|\phi\|^2 \leq \frac{(b-a)^2}{2}\|\phi'\|^2.$$

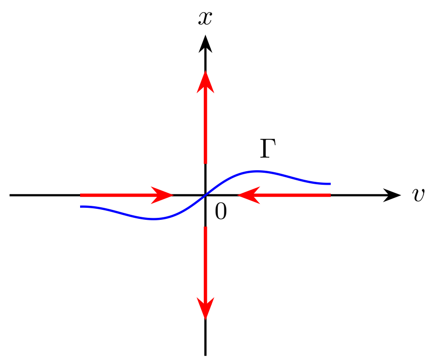


FIGURE 2. Condition $\mathcal{H}^* \geq 0$ is not necessary. For the Lagrangian $L(t,x,v) = v^2 - x^2$, the curve $x_* = 0$ is a saddle point of $\mathcal{H}^*$ but still a local minimizer.

The Hessian matrix $\mathcal{H}$ is defined with respect to the variables (v, x), where v and x are treated as independent. However, in the second-order variation, these variables are coupled through the constraint $v = x'$, which defines a lower-dimensional manifold $\Gamma = \{(v,x) \mid v = x'\}$. The functional $I(x)$ attains a local minimum on Γ, not over the full space of (v, x). Moreover, the variation involves integration over the interval $[a, b]$, which averages out pointwise contributions.

3.3. **Hessian of the functional.** As we shall demonstrate shortly, the derivative term x' governs the scaling and can dominate the contribution of the Hessian in the second-order variation. Consequently, in Example 3.1, even though $\mathcal{H}$ is indefinite and only positive definite in the v-direction, it is still possible for x_* to be a local minimizer.

We now express the second-order variation in an operator form. Assume that $L \in C^3$. By applying integration by parts, the cross term satisfies

$$
\begin{aligned}
&(B\phi, \phi') = -((B\phi)', \phi) = -(B'\phi, \phi) - (B\phi', \phi),\\
\implies \quad & 2(B\phi, \phi') = -(B'\phi, \phi).
\end{aligned}
$$

Therefore, the second-order variation becomes

$$\delta^2 I(x, \phi) = (S\phi, \phi)_1 := (P(t)\phi', \phi') + (Q(t)\phi, \phi),$$

where

$$
\begin{aligned}
&P(t) = A(t) = L_{vv}(t, x(t), x'(t)),\\
&Q(t) = C(t) - B'(t) = L_{xx}(t, x(t), x'(t)) - \frac{\mathrm{d}}{\mathrm{d}t} L_{xv}(t, x(t), x'(t)).
\end{aligned}
$$

The operator $S : H_0^1(a,b) \to H_0^1(a,b)$ is defined as: for $\phi \in H_0^1(a,b)$, $S\phi \in H_0^1(a,b)$ s.t.

$$(S\phi, \psi)_1 = (P(t)\phi', \psi') + (Q(t)\phi, \psi), \quad \forall \psi \in H_0^1(a,b). \tag{12}$$

As the right hand side of (12) with a given ϕ is a linear continuous functional on $H_0^1(a,b)$ (treating ψ as the variable), by the Riesz representation theorem, the operator S is well defined. It is obviously self-adjoint, i.e., $(S\phi, \psi)_1 = (\phi, S\psi)_1$.

A sufficient condition for x_* to be a local minimizer is

$$\delta^2 I(x_*, \phi) > 0 \quad \forall \phi \in H_0^1(a,b) \qquad \iff \qquad (S^*\phi, \phi)_1 > 0 \quad \forall \phi \in H_0^1(a,b),$$

that is, S^* is self-adjoint and positive definite (SPD) on $H_0^1(a,b)$. A necessary condition for local minimality is then $S^* \geq 0$.

In summary, by applying integration by parts, the Hessian matrix $\mathcal{H}$ of the Lagrangian is transformed into a diagonal operator S. Although these operators act on different spaces, their actions are equivalent when restricted to the manifold $\Gamma = \{(\phi', \phi)\}$. Since $\delta^2 I(x_*, \phi) = (S^*\phi, \phi)_1$ for all $\phi \in H_0^1(a,b)$, the operator S^* effectively plays the role of the Hessian in the variational setting.

3.4. **Legendre condition.** In the Hessian matrix $\mathcal{H}$, the component $A = L_{vv}$ dominates in the sense that it acts on the derivatives of the test function: $(A\phi', \phi')$. Using a scaling argument, one can construct a localized test function ϕ supported near a particular point $t \in [a,b]$, with support length h and height 1—for example, a hat function in $H_0^1(a,b)$. For such a function, we have the scaling estimates

$$(\phi', \phi') = \mathcal{O}(h^{-1}), \quad (\phi', \phi) = \mathcal{O}(h), \quad \text{and} \quad (\phi, \phi) = \mathcal{O}(h^2).$$

If we wish to use C^1 test functions instead of those merely in H^1, we may apply a mollifier to the hat function to obtain the desired smoothness while preserving the essential scaling behavior.

Legendre Condition

If $x_* \in \mathcal{M}$ is a local minimizer of $I(\cdot)$. Then the following Legendre condition holds

$$P^*(t) = A^*(t) := L_{vv}(t, x_*(t), x_*'(t)) \geq 0, \quad \forall t \in [a,b]. \tag{13}$$

Proof. Construct a hat function $\phi \in H_0^1(a,b)$ with support $[t-h, t+h]$ and height h. Then $|\phi'|^2 = 1$ only in $[t-h, t+h]$ and zero otherwise. From the second variation condition $\delta^2 I(x, \phi) \geq 0$, we conclude

$$\int_{t-h}^{t+h} \left[P^*(t) + |Q^*(t)|h^2\right] \, \mathrm{d}t \geq 0.$$

As h and t are arbitrary, we conclude the Legendre condition. □

Although the strict (or strengthened) Legendre condition $A^* = P^* > 0$ ensures that the leading-order part of the Hessian is positive definite, it is not sufficient for local minimality, since A^* constitutes only a portion—albeit the dominant portion—of $\mathcal{H}^*$ or S^*. Readers are encouraged to construct explicit counterexamples illustrating this point.

Instead, one should examine the full operator S^*, as the second variation satisfies $\delta^2 I(x_*, \phi) = (S^*\phi, \phi)_1$. Thus, S^* plays the role of the Hessian of the functional in the calculus of variations framework. Conditions involving $\mathcal{H}^*$, A^*, and S^* are summarized as follows:

$$\mathcal{H}^* > 0 \Longrightarrow x_* \text{ is a local minimizer } \not\Longleftarrow \mathcal{H}^* \geq 0.$$
$$A^* > 0 \not\Longleftarrow x_* \text{ is a local minimizer } \Longrightarrow A^* \geq 0.$$
$$S^* > 0 \Longrightarrow x_* \text{ is a local minimizer } \Longrightarrow S^* \geq 0.$$

In the following, we will focus on conditions on $S^* \geq 0$ and $S^* > 0$.

3.5. **Conjugate points.** Assuming $P^* > 0$ and $Q^* \geq 0$, which is roughly equivalent to the strong sufficient condition $\mathcal{H}^* > 0$, we can conclude that $S^* > 0$. However, the converse is not necessarily true: when $S^* > 0$, we cannot guarantee that $Q^*(t) \geq 0$ for all $t \in [a, b]$. It is possible for Q^* to take negative values while the overall operator S^* remains positive definite. See Example 3.1.

To study the condition $S^* > 0$ more deeply, we introduce the concept of *conjugate points*. We begin with an example in which the notion of a conjugate point has a clear geometric interpretation.

Exercise 3.3. Consider two points a and b on the unit sphere S^2. The geodesic connecting a and b lies along a great circle, and there are typically two such arcs connecting the points. Which one is the shortest?

Let $\hat{a} = -a$ denote the point on the sphere antipodal to a, known as the *conjugate point* of a. Show that the shortest geodesic from a to b is the one that does not pass through the conjugate point $\hat{a}$.

Hint: the computation of the geodesic can be simplified by selecting the plane containing the points $(a, 0, b)$ to be the x-y plane, and choosing $a = (1, 0, 0)$ without loss of generality. □

In the general case, the concept of a conjugate point is not so straightforward. Given a number $\sigma \in (a, b]$, we can embed the subspace $H_0^1(a, \sigma) \hookrightarrow H_0^1(a, b)$ by the zero extension, and naturally restrict the operator S to $H_0^1(a, \sigma)$. More precisely,

$$(S_\sigma \phi, \psi)_1 := (P(t)\phi', \psi') + (Q(t)\phi, \psi), \qquad \phi, \psi \in H_0^1(a, \sigma).$$

The notation S_σ is used to indicate the dependence on the parameter σ.

Definition 3.4 (Conjugate points)**.** *Let x_* be a solution to the Euler-Lagrange equation. A conjugate point of a (along x_*) is a number $\sigma \in (a, b]$ such that S_σ^* has a zero eigenvalue on $H_0^1(a, \sigma)$. In other words, there exists a nonzero function $u \in H_0^1(a, \sigma)$ satisfying $S_\sigma^* u = 0$, with the following strong form:*

$$-\left(P^*(t)u'(t)\right)' + Q^*(t)u(t) = 0, \quad t \in (a, \sigma), \quad u(a) = u(\sigma) = 0. \tag{14}$$

If a conjugate point exists, we can use the eigenfunction u from (14) as the variation. This leads to $\delta^2 I(x_*, u) = 0$, which means we cannot conclude whether x_* is a local minimizer.

When $P^* > 0$, the operator S^* becomes a Sturm–Liouville operator, which is compact and thus has a countable set of real eigenvalues. More specifically, for σ sufficiently close to a, we may approximate the coefficients $P(t)$ and $Q(t)$ by constants: set $p = P(a) > 0$ and $q = Q(a)$. This yields the linear eigenvalue problem

$$-pu'' + qu = \lambda u, \quad u(a) = u(\sigma) = 0.$$

The eigenfunctions are given by $u_k = \sin\left(\frac{k\pi(x-a)}{\sigma - a}\right)$ for $k \in \mathbb{N}$, with corresponding eigenvalues

$$\lambda_k = q + \frac{pk^2\pi^2}{(\sigma - a)^2}.$$

In particular, for $k = 1$, we obtain, even for $q < 0$,

$$\lambda_{\min}(S_\sigma) = q + \frac{p\pi^2}{(\sigma - a)^2} > 0, \quad \text{for } \sigma - a \ll 1. \tag{15}$$

This provides information about the behavior of the smallest eigenvalue of S_σ as a function of σ. We will prove the following result.

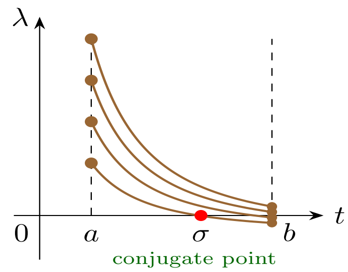


FIGURE 3. Eigenvalues of S_σ are decreasing with respect to σ. At the conjugate point σ, $\lambda_{\min}(S_\sigma^*) = 0$ and the corresponding eigen-function u satisfies $\delta^2 I(x_*, u) = 0$.

Proposition 3.5. *Assume $P(t) > 0$ for all $t \in [a, b]$. Then $\lambda_{\min}(S_\sigma)$ is a decreasing function of $\sigma \in (a, b]$.*

Proof. Intuitively, (15) shows that locally $1/(\sigma - a)^2$ is a decreasing function of σ, and the minimal frequency corresponds to $k = 1$. The associated eigenfunction is a sine function. But the closed form (15) of eigenvalues only holds approximately for σ near a.

For the self-adjoint operator S_σ, the smallest eigenvalue can be characterized as

$$\lambda_{\min}(S_\sigma) = \inf_{v \in H_0^1(a,\sigma)} \frac{(S_\sigma v, v)_1}{(v, v)}. \tag{16}$$

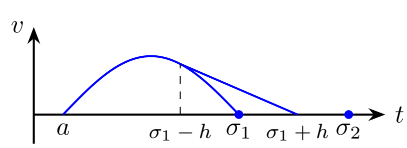
Take $\sigma_1 < \sigma_2$. Then $H_0^1(a, \sigma_1) \subset H_0^1(a, \sigma_2)$. From (16), we immediately have

$$\lambda_{\min}(S_{\sigma_1}) \geq \lambda_{\min}(S_{\sigma_2}).$$

Next we show that the inequality is strict. Let v be the eigenfunction corresponding to $\lambda_{\min}(S_{\sigma_1})$. If $v'(\sigma_1) = 0$, then together with the boundary condition $v(\sigma_1) = 0$, it follows that v must vanish identically near σ_1. In this case, we may shift σ_1 slightly to the left so that $v'(\sigma_1) \neq 0$. Without loss of generality, assume $v'(\sigma_1) < 0$.

We now construct a function $\tilde{v} \in H_0^1(a, \sigma_1 + h)$ and show that

$$\frac{(S_{\sigma_2} \tilde{v}, \tilde{v})_1}{(\tilde{v}, \tilde{v})} < \frac{(S_{\sigma_1} v, v)_1}{(v, v)} \quad \text{for sufficiently small } h > 0.$$



Near $t = \sigma_1$, we approximate the coefficients $P(t)$ and $Q(t)$ by constants $P(\sigma_1)$ and $Q(\sigma_1)$, and v' by the constant $v'(\sigma_1)$. By construction, the slope of the extension is halved, i.e., $\tilde{v}' = \frac{1}{2} v'(\sigma_1)$ in the extended region.

By the scaling argument, the reduction in the derivative term will dominate the increase in the L^2 term. In fact, one can show

$$\|\tilde{v}'\|^2 = \|v'\|^2 - \mathcal{O}(h), \quad \text{and} \quad \|\tilde{v}\|^2 = \|v\|^2 + \mathcal{O}(h^3).$$

Therefore, for $h \ll 1$, the Rayleigh quotient strictly decreases, implying

$$\lambda_{\min}(S_{\sigma_1}) > \lambda_{\min}(S_{\sigma_2}),$$

as claimed. □

Proposition 3.5 can be generalized to every eigenvalue of S_σ and the full proof can be found in [3].

3.6. **Jacobi condition.** The operator S^* plays the role of Hessian.

Necessary condition to be a local minimum: no interior conjugate point

Suppose $P^* := L_{vv}(t, x_*(t), x_*'(t)) > 0$ for all $t \in [a, b]$. If x_* is a local minimizer of $I(\cdot)$, then

(1) x_* satisfies the E-L equation;
(2) there is no interior conjugate point in (a, b).

Note that since $\delta^2 I(x_*, \phi) \geq 0$ for any admissible variation ϕ at a local minimizer x_*, the right endpoint b could potentially be a conjugate point. However, the necessary condition discussed above implies that there must be no interior conjugate point in (a, b). Otherwise, by Proposition 3.5, the existence of a conjugate point $\sigma_1 \in (a, b)$ implies that for any $\sigma_2 \in (\sigma_1, b]$, the operator $S^*_{\sigma_2}$ possesses a negative eigenvalue. In particular, for the corresponding eigenfunction u, we would have $\delta^2 I(x_*, u) < 0$, contradicting the assumption that x_* is a local minimizer. An elementary proof using regularity results will be provided in Exercise 3.7.

We present the following sufficient conditions for a local minimizer.

Sufficient conditions: E-L + strict Legendre condition + Jacobi condition

Suppose

(1) x_* satisfies the E-L equation of $I(\cdot)$;
(2) $P^*(t) := L_{vv}(t, x_*(t), x_*'(t)) > 0$ for all $t \in [a, b]$;
(3) no conjugate point in $(a, b]$,

then x_* is a strict local minimizer.

As $\lambda_{\min}(S^*_\sigma) > 0$ for σ sufficiently close to a and $\lambda_{\min}(S^*_\sigma)$ is a decreasing function of σ, the "no conjugate point condition in $(a, b]$" implies $\lambda_{\min}(S^*) = \lambda_0 > 0$. Consequently,

$$\delta^2 I(x_*, \phi) = (S^*\phi, \phi)_1 \geq \lambda_0 \|\phi\|^2, \quad \forall \phi \in H_0^1(a, b).$$

To strengthen this inequality to the H^1-norm, we use the following result.

Exercise 3.6. Assume $P^* > 0$ for all $t \in [a, b]$ and that

$$\delta^2 I(x_*, \phi) \geq \mu \|\phi\|^2 \quad \text{for all } \phi \in H_0^1(a, b).$$

Then the second variation is also coercive in the H^1 norm, i.e., there exists a constant $\lambda > 0$ such that

$$\delta^2 I(x_*, \phi) \geq \lambda \|\phi\|_1^2 \quad \forall \phi \in H_0^1(a, b),$$

where $\|\phi\|_1^2 = \|\phi\|^2 + \|\phi'\|^2$ denotes the H^1-norm.

How to check whether there is a conjugate point? Solve the Jacobi equation

$$-(P^*(t)u'(t))' + Q^*(t)u(t) = 0, \qquad u(a) = 0, u'(a) = 1. \tag{17}$$

And check whether u has a zero in the open interval (a, b) or $(a, b]$.

Exercise 3.7. Let x_* be an extremal which satisfies the strengthened Legendre condition. We shall prove the necessary condition "no interior conjugate point":

If x_* is a weak local minimizer, then there is no conjugate point to a in the interval (a, b).

Suppose a conjugate point $\sigma \in (a, b)$ exists. By definition of the conjugate point, there exists a non-trivial solution u of Jacobi's equation

$$-(P^*u')' + Q^*u = 0, \ t \in (a, \sigma), \quad u(a) = u(\sigma) = 0.$$

(1) Prove that $u'(\sigma) \neq 0$ and that

$$\int_a^\sigma \left[P^*(t)u'(t)^2 + Q^*(t)u(t)^2\right] \, \mathrm{d}t = 0.$$

(2) Extend u to $[a, b]$ by zero. Consider the minimization problem

$$\min_{y \in \mathrm{Lip}_0[a,b]} I(y) := \int_a^\sigma \left[P^*(t)y'(t)^2 + Q^*(t)y(t)^2\right] \, \mathrm{d}t.$$

Then u is a minimizer. Use the regularity result (Theorem 2.2) at the conjugate point σ to obtain a contradiction.

4. Strong Minima

We discuss perturbations in different norms. For the function space $\phi \in W_0^{1,\infty}(a, b)$, the default norm is

$$\|\phi\|_{1,\infty,[a,b]} = \max\{\|\phi\|_{\infty,[a,b]}, \|\phi'\|_{\infty,[a,b]}\}.$$

To simplify notation, we shall omit the interval in the norm. As a subspace of $L^\infty[a, b]$, a weaker norm $\|\cdot\|_\infty$ can be used. Using different norms, we define different open balls around a point:

$$B(x_*, \epsilon) = \{x \in C[a, b] : \|x - x_*\|_\infty < \epsilon\},$$
$$B_1(x_*, \epsilon) = \{x \in C_0^1[a, b] : \|x - x_*\|_{1,\infty} < \epsilon\}.$$

These two norms satisfies

$$\|\phi\|_\infty \le \|\phi\|_{1,\infty} \Longrightarrow B_1(x_*, \epsilon) \subseteq B(x_*, \epsilon).$$

Thus, the topology induced by a stronger norm admits more open sets. For instance, both $\|\cdot\|_\infty$ and $\|\cdot\|_{1,\infty}$ topologies include the strip as an open set, but only the latter includes the square. Hence, $\|\cdot\|_{1,\infty}$ is stronger than $\|\cdot\|_\infty$.

Definition 4.1. *A function $x_* \in \mathcal{M}$ is a* strong local minimizer *of $I(\cdot)$ if there exists $\epsilon > 0$ such that*

$$I(x_* + \phi) \ge I(x_*), \quad \forall \phi \in H_0^1(a, b) \cap B(x_*, \epsilon). \tag{18}$$

If the condition is changed to $\phi \in B_1(x_*, \epsilon)$, then it is called a weak local minimizer.

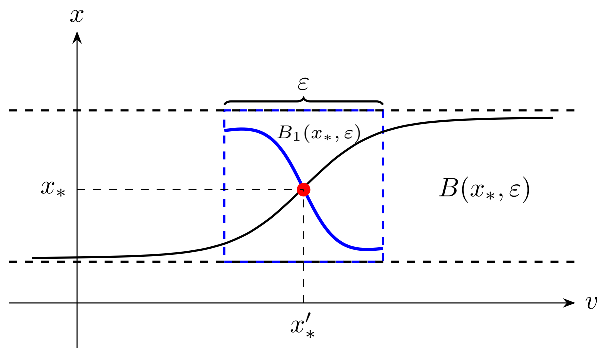


FIGURE 4. Perturbation in the weak minimum vs. the strong minimum.

Figure 4 illustrates the geometry of perturbations. The small square represents the $\|\cdot\|_\infty$ ball of radius ϵ centered at the extremum (x'_*, x_*) in the (v, x)-plane. Since $v = x'$, the variables (v, x) lie on a curve passing through this point.

For a weak minimum, the perturbation is restricted to this curve within the square, with both $\|\phi\|_\infty$ and $\|\phi'\|_\infty$ bounded by ϵ. For a strong minimum, only $\|\phi\|_\infty \leq \epsilon$ is required; $\|\phi'\|_\infty$ may be unbounded. As a result, the curve may extend beyond the square but remains within the horizontal strip $B_\epsilon(x_*, \|\cdot\|_\infty)$.

- *Strong* minimum: perturbation lies in a *larger* ϵ-ball of a *weaker* norm.
- *Weak* minimum: perturbation lies in a *smaller* ϵ-ball of a *stronger* norm.

Obviously, a strong minimum is also a weak minimum but not vice verse. An example can be easily constructed by using a Lagrangian of x' only. The variation ϕ is small in L^2 but not in H^1. For example, a high frequency perturbation $\phi = \epsilon \sin(k\pi t)$.

Example 4.2 (Example 14.14 (page 301 and 318 in [2]))**.** Consider

$$\min \int_0^1 (x'(t))^3 \,\mathrm{d}t : x(0) = 0, x(1) = 1.$$

A weak local minimum is $x_*(t) = t$ satisfying E-L condition, strengthened Lendre condition, no conjugate condition, but is not a strong local minimum.

All minima discussed previously are *weak local minima*. The sufficient condition

$$\delta^2 I(x_*, \phi) \geq \lambda \|\phi\|_1^2$$

for x_* to be a weak local minimizer is no longer sufficient to guarantee that x_* is a *strong* local minimizer.

Consider the simple example from calculus: minimize a smooth function $f(x)$. Suppose 0 is a critical point. Then we have the Taylor expansion

$$f(\phi) - f(0) = \frac{1}{2} f''(0)\phi^2 + \frac{1}{6} f'''(\xi)\phi^3$$

for some ξ between 0 and ϕ. If $f''(0) \geq \lambda > 0$, then for $|\phi|$ sufficiently small, the third-order term satisfies

$$|f'''(\xi)||\phi|^3 \leq \lambda |\phi|^2,$$

so that

$$f(\phi) - f(0) \geq \frac{\lambda}{3} |\phi|^2 > 0,$$

which implies that 0 is a local minimum.

In the variational setting, the situation is more subtle due to the presence of different norms. For instance, if we consider a perturbation ϕ satisfying $\|\phi\|_\infty < \varepsilon$, we no longer have any control over $\|\phi'\|_\infty$. In the expansion

$$I(x) - I(x_*) = \delta^2 I(x_*, \phi) + \text{Remainder}(\phi, \phi'),$$

the remainder typically depends on both $\|\phi\|_\infty$ and $\|\phi'\|_\infty$. Since smallness of $\|\phi\|_\infty$ does not imply smallness of $\|\phi'\|_\infty$, the quadratic lower bound on $\delta^2 I(x_*, \phi)$ does not suffice to ensure that $I(x) \geq I(x_*)$ for all ϕ with small $\|\phi\|_\infty$. In other words, the condition for a weak local minimum does not imply a strong local minimum.

We need additional conditions to control the derivative norm outside the small ball $B_1(x_*, \epsilon)$. Such a condition should be imposed to account for arbitrary perturbations in the v-direction.

Given a C^1 Lagrangian $L(t, x, v)$, define the *Weierstrass excess function* as

$$\mathfrak{E}_L(t, x, v, q) = L(t, x, q) - L(t, x, v) - (q - v) \cdot L_v(t, x, v). \tag{19}$$

This represents the deviation from the linear approximation of $L(t, x, \cdot)$ at the point v. If $\mathfrak{E}_L(t, x, v, q) \geq 0$ for all (v, q), then $L(t, x, v)$ is convex with respect to v, and $\mathfrak{E}_L$ is known as the *Bregman divergence*.

A necessary condition for a strong minimum requires the positivity of $\mathfrak{E}_L$ only at the minimizer. That is, the function $v \mapsto L(t, x, v)$ does not need to be convex globally; see Fig. 4, but it must lie above the tangent line at v_*.

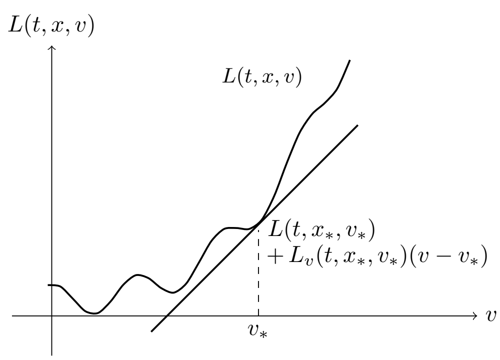


FIGURE 5. Weierstrass condition.

Weierstrass condition

If $x_* \in \mathcal{M}$ is a local strong minimizer of $I(\cdot)$. Then the following Weierstrass condition holds

$$\mathfrak{E}_L(t, x_*, x_*'(t), x_*'(t) + \xi) \geq 0, \quad \forall \xi \in \mathbb{R}^N, \forall t \in [a, b]. \tag{20}$$

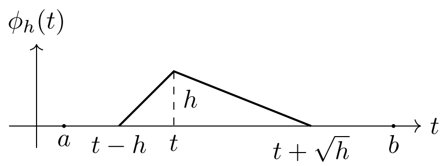
Proof. We outline a proof of the Weierstrass condition as given in [1, page 45].

We begin with the definition of a local minimizer:

$$I(x_* + \phi) - I(x_*) - \delta I(x_*, \phi) \geq 0, \quad \forall \phi \in C_0^1(a, b). \tag{21}$$

We include the vanishing first variation $\delta I(x_*, \phi)$ so that the integrand becomes

$$\int_a^b \left[L(t, x_* + \phi, x_*' + \phi') - L(t, x_*, x_*') - L_x \cdot \phi - L_v \cdot \phi'\right] \mathrm{d}t \geq 0.$$



Given a fixed vector $\xi \in \mathbb{R}^N$, we construct a perturbation ϕ_h as shown in Fig. 4, and use $\phi = \xi\phi_h$ as the variation in (21). The support of ϕ_h is limited to the intervals $(t-h, t)$ and $(t, t+\sqrt{h})$, so only these contribute to the integral. By the scaling argument, the integral over $(t, t+\sqrt{h})$ is $o(\xi h)$, while the integral over $(t-h, t)$ is $O(h)$. Thus, we focus on the limit:

$$\begin{aligned}&\lim_{h\to 0} \frac{1}{h} \int_{t-h}^{t} \left[L(t, x_* + \phi, x_*' + \phi') - L(t, x_*, x_*') - L_x \cdot \phi - L_v \cdot \phi'\right] \mathrm{d}t \\ &= L(t, x_*(t), x_*'(t) + \xi) - L(t, x_*(t), x_*'(t)) - \xi \cdot L_v(t, x_*(t), x_*'(t)),\end{aligned}$$

which leads to the Weierstrass condition (20).

The slope of ϕ_h from the left is 1 and from the right is $-\sqrt{h}$ so that when $h \to 0$, $\xi\phi_h' \to \xi$ not $-\xi$. □

A sufficient condition for an extremal to be a strong minimizer will be established using extremal fields, which will be discussed later in the context of Hamilton–Jacobi theory. Roughly speaking, this condition requires that the function $v \mapsto L(t, x, v)$ be locally convex in a neighborhood of v_*.

References

[1] K.-C. Chang. *Lecture Notes on Calculus of Variations*, volume 6. World Scientific, 2016. 15

[2] F. Clarke. *Functional analysis, calculus of variations and optimal control*, volume 264. Springer Science & Business Media, 2013. 3, 14

[3] R. S. Manning. Conjugate Points Revisited and Neumann–Neumann Problems. *SIAM Review*, 51(1):193–212, Feb. 2009. 12