

Appendix C

Legendre Transforms, Calculus of Variations, and Mechanics Principles

C.1 Legendre Transforms

Legendre transforms map functions in a vector space to functions in the dual space. From a theoretical perspective, they play a fundamental role in the construction of dual Banach spaces in functional analysis and the concepts of tangential coordinates and projective duality in algebraic geometry. Within the realm of model development for smart systems, Legendre transforms are employed when defining thermodynamic potentials and establishing the correspondence between Lagrangian and Hamiltonian frameworks for dynamic systems.

Definitions

A function f is termed *convex* if

$$f(\alpha x + (1 - \alpha)y) \leq \alpha f(x) + (1 - \alpha)f(y)$$

for all x and y within the domain and all α , $0 < \alpha < 1$ — see Figure C.1. The definition is geometric and does not require that f be differentiable. If sufficiently smooth, however, f will be convex if and only if $f''(x) \geq 0$.

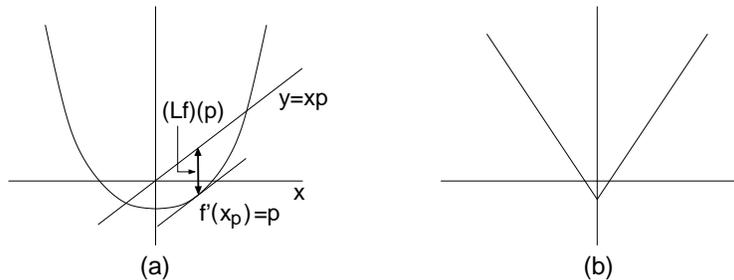


Figure C.1. Convex functions f which are (a) differentiable throughout the domain and (b) nondifferentiable at points in the domain.

438 Appendix C. Legendre Transforms, Calculus of Variations, Mechanics Principles

Let $f : \mathbb{R}^1 \rightarrow \mathbb{R}^1$ be a convex function. The Legendre transform $g(p) = (Lf)(p)$ is defined by

$$(Lf)(p) = \sup_x [xp - f(x)].$$

Note that this specifies $x = g(p)$ as a function of p . If f is differentiable, the Legendre transform can be expressed as

$$\begin{aligned} (Lf)(p) &= \max_x [xp - f(x)] \\ &= x_p p - f(x_p) \end{aligned}$$

where x_p solves

$$p = f'(x_p). \tag{C.1}$$

The existence and uniqueness of x_p in this case is due to the convexity of f .

If we employ a dot product rather than the scalar product, the analogous definition

$$(Lf)(p) = \sup_x [x \cdot p - f(x)] \tag{C.2}$$

holds for convex $f : \mathbb{R}^n \rightarrow \mathbb{R}^1$.

Properties

Two properties of the Legendre transform are of fundamental importance for both analysis and applications: (i) it maps convex functions to convex functions, and (ii) the Legendre transform is self-dual or an involution. The first plays a fundamental role in the optimization of energy relations since it dictates the manner through which minimum energy states for one energy definition are related to those of its transform. The second property states that $L(Lf) = f$. For differentiable functions, this follows from the property that if p and x are related by $p = f'(x)$, then $x = g'(p)$. Proofs and ramifications of these properties can be found in [15].

Further attributes of the Legendre transform in 1-D are illustrated by the following examples.

Example 1. Let $f(x) = x^2$. From the condition $p = 2x$, it follows that

$$g(p) = (Lf)(p) = \frac{p^2}{2} - \frac{p^2}{4} = \frac{p^2}{4}.$$

Example 2. Let $f(v) = \frac{1}{2}mv^2$ denote the kinetic energy for a particle of mass m . The necessary condition (C.1) yields the momentum equation

$$p = mv.$$

The Legendre transform in this case is

$$g(p) = \frac{p^2}{2m}.$$

Further ramifications of this relation for Hamiltonian mechanics formulations will be provided in Section C.3.

Example 3. Consider the elastic Helmholtz energy relation

$$\psi(\varepsilon) = \frac{1}{2}Y\varepsilon^2$$

which results from (2.18) when polarization is neglected and strains ε are restricted to 1-D. The negative Legendre transform is

$$\begin{aligned} -(L\psi)(\sigma) &= -\left[\frac{\sigma^2}{Y} - \frac{1}{2}\left(\frac{\sigma}{Y}\right)^2\right] \\ &= -\frac{1}{2}s\sigma^2 \end{aligned}$$

where σ is an applied stress and $s = \frac{1}{Y}$ is the 1-D compliance. Comparison with (2.22) illustrates that

$$G(\sigma) = -(L\psi)(\sigma)$$

defines the elastic Gibbs energy for the system. Similarly, it is shown in Section 2.2.3 that $G(E) = -(L\psi)(E)$ for the Helmholtz energy $\psi(P) = \frac{1}{2}\alpha P^2$.

C.2 Principles from the Calculus of Variations

To provide fundamental relations used when establishing the Lagrange mechanics framework in Section C.3, we summarize selected principles pertaining to the calculus of variations. Additional details can be found in [15, 307, 505].

Gateaux and Fréchet Differentials

Calculus of variations is concerned with the extrema of functionals so we begin with a summary of differential theory for vector spaces. Throughout this discussion, X is a vector space, Y is a normed space, and $T : D \subset X \rightarrow Y$ is a (possibly nonlinear) transformation. For the case $Y = \mathbb{R}$, the transformation is a real-valued functional which we will denote by J .

Definition C.2.1. Consider $x \in D \subset X$ and arbitrary $\eta \in X$. If the limit

$$\delta T(x; \eta) = \lim_{\epsilon \rightarrow 0} \frac{1}{\epsilon} [T(x + \epsilon\eta) - T(x)]$$

exists for each $\eta \in X$, T is said to be Gateaux differentiable at x and $\delta T(x; \eta)$ is termed the Gateaux differential of T at x with increment or perturbation η .

For functionals J , the Gateaux differential, when it exists, is

$$\delta J(x; \eta) = \left. \frac{d}{d\epsilon} J(x + \epsilon\eta) \right|_{\epsilon=0}.$$

Note that for each fixed $x \in D$, $\delta J(x; \eta)$ is a functional with respect to $\eta \in X$.

Definition C.2.2. T is said to be Fréchet differentiable at $x \in D$ in the normed space X if for each $\eta \in X$, there exists $\delta T(x; \eta) \in Y$ which is linear, continuous with respect to η , and satisfies

$$\lim_{\|\eta\| \rightarrow 0} \frac{\|T(x + \eta) - T(x) - \delta T(x; \eta)\|}{\|\eta\|} = 0.$$

When it exists, $\delta T(x; \eta)$ is termed the Fréchet differential of T at x with increment η .

The Gateaux differential generalizes the concept of directional derivatives whereas the Fréchet differential generalizes the definition of differentiability. Because the Gateaux differential requires no norm on X , it cannot be directly used to establish continuity and is significantly weaker than the Fréchet differential. This is illustrated by the calculus example

$$f(x, y) = \begin{cases} 0 & , (x, y) \in \Omega_1 = \{(x, y) \mid y \neq x^2 \text{ or } x = y = 0\} \\ 1 & , (x, y) \in \Omega_2 = \{(x, y) \mid y = x^2 \text{ and } x \neq 0, y \neq 0\}. \end{cases}$$

All directional derivatives exist at $(0, 0)$ but the function is both discontinuous and nondifferentiable at that point.

We note that the existence of the Fréchet differential implies the existence of the Gateaux differential in which case the two will be equal.

Extrema of Functionals

The following theorem establishes a necessary condition for a functional to have an extremum (minimum or maximum) at the point x_0 .

Theorem C.1. Let the functional $J : X \rightarrow \mathbb{R}$ have a Gateaux differential $\delta J(x; \eta)$. If J has an extremum at x_0 , then $\delta J(x_0; \eta) = 0$ for all $\eta \in X$.

Proof. If J has an extremum at x_0 , it follows that the function $J(x_0 + \epsilon\eta)$ of the real variable ϵ has an extremum at $\epsilon = 0$. This implies that

$$\left. \frac{d}{d\epsilon} J(x_0 + \epsilon\eta) \right|_{\epsilon=0} = 0$$

and hence $\delta J(x_0; \eta) = 0$. □

Points x_0 at which extrema occur are termed *stationary points*. We note that in the context of mechanics, Theorem C.1 is often referred to as *Hamilton's principle* or *Hamilton's principle of least motion*.

Euler–Lagrange Equations

Consider the problem of finding a function x which minimizes the functional

$$J = \int_{t_0}^{t_1} \mathcal{L}[x(t), \dot{x}(t), t] dt. \tag{C.3}$$

The function \mathcal{L} is assumed to be continuous in x, \dot{x}, t and have continuous partial derivatives with respect to x and \dot{x} . We also assume that the endpoints $x(t_0)$ and $x(t_1)$ are fixed.

To specify the admissible class of solutions, we consider variations of the form

$$\hat{x}(t) = x(t) + \epsilon\eta(t)$$

where η satisfies

$$\begin{aligned} (i) \quad & \eta \in C^1[t_0, t_1] \\ (ii) \quad & \eta(t_0) = \eta(t_1) = 0. \end{aligned} \tag{C.4}$$

The first criterion guarantees the continuity of solutions and their temporal derivatives whereas the second guarantees that

$$\hat{x}(t_0) = x(t_0) \quad , \quad \hat{x}(t_1) = x(t_1)$$

in accordance with the condition of fixed endpoints. The second condition is depicted in Figure C.2.

The Gateaux differential is

$$\begin{aligned} \delta J(x; \eta) &= \left. \frac{d}{d\epsilon} \int_{t_0}^{t_1} \mathcal{L}(x + \epsilon\eta, \dot{x} + \epsilon\dot{\eta}, t) dt \right|_{\epsilon=0} \\ &= \int_{t_0}^{t_1} \frac{\partial \mathcal{L}}{\partial x}(x, \dot{x}, t)\eta(t) dt + \int_{t_0}^{t_1} \frac{\partial \mathcal{L}}{\partial \dot{x}}(x, \dot{x}, t)\dot{\eta}(t) dt \end{aligned}$$

which can be verified to be a Fréchet differential. Under the assumption that $\frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \dot{x}}$ exists and is continuous, integration by parts and application of Theorem C.1 yields

$$\begin{aligned} \delta J(x; \eta) &= \int_{t_0}^{t_1} \left[\frac{\partial \mathcal{L}}{\partial x}(x, \dot{x}, t) - \frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \dot{x}}(x, \dot{x}, t) \right] \eta(t) dt + \left. \frac{\partial \mathcal{L}}{\partial \dot{x}}(x, \dot{x}, t)\eta(t) \right|_{t_0}^{t_1} \\ &= \int_{t_0}^{t_1} \left[\frac{\partial \mathcal{L}}{\partial x}(x, \dot{x}, t) - \frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \dot{x}}(x, \dot{x}, t) \right] \eta(t) dt \\ &= 0 \end{aligned}$$

which must hold for all η satisfying the admissibility conditions (C.4). Because η is continuous, it follows that the optimal x must satisfy the Euler–Lagrange equation

$$\frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \dot{x}}(x, \dot{x}, t) - \frac{\partial \mathcal{L}}{\partial x}(x, \dot{x}, t) = 0. \tag{C.5}$$

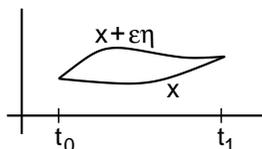


Figure C.2. Admissible variations in the trajectory x .

442 Appendix C. Legendre Transforms, Calculus of Variations, Mechanics Principles

A derivation of (C.5) which relaxes the *a priori* continuity condition on $\frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \dot{x}}$ is provided in [307].

Example 4. Let $x = \mathbb{R}$ and take $\mathcal{L}(x, \dot{x}, t) = \sqrt{1 + \dot{x}^2}$ so that

$$J = \int_{t_0}^{t_1} \sqrt{1 + \dot{x}^2} dt$$

defines the length of the curve between the endpoints $x(t_0)$ and $x(t_1)$. Here $\frac{\partial \mathcal{L}}{\partial x} = 0$ and $\frac{\partial \mathcal{L}}{\partial \dot{x}} = \frac{\dot{x}}{\sqrt{1 + \dot{x}^2}}$ so the Euler–Lagrange equation is

$$\frac{d}{dt} \left(\frac{\dot{x}}{\sqrt{1 + \dot{x}^2}} \right) = 0.$$

Integration yields $\dot{x} = c_1$ and hence $x(t) = c_1 t + c_2$. As expected, the length is minimized by a straight line with c_1 and c_2 determined by $x(t_0)$ and $x(t_1)$.

C.3 Classical, Lagrangian and Hamiltonian Mechanics

We summarize here basic tenets of classical, Lagrangian and Hamiltonian mechanics to provide aspects of the framework employed when constructing constitutive and structural models for smart material systems.

Classical Mechanics

Classical or Newtonian mechanics can be described as the physics of forces or moments acting on a body. To simplify the discussion, we consider only forces acting on a point particle of mass m and refer the reader to [15] for discussion regarding more complex systems. We let $\mathbf{r} = \mathbf{r}(x_1, x_2, x_3, t)$ denote the position of the particle and $\mathbf{v} = \dot{\mathbf{r}}$ denote its velocity.

One of the cornerstones of classical mechanics is Newton’s second law

$$\mathbf{F} = \frac{d}{dt}(m\mathbf{v}) \tag{C.6}$$

which states that the change in momentum $\mathbf{p} = m\mathbf{v}$ is equal to the sum of all applied forces. When the mass is time invariant, this yields the familiar relation

$$\mathbf{F} = m\mathbf{a}.$$

Three scalar quantities which are fundamental for quantifying the static and dynamic response of the body are the work, kinetic energy and potential energy. The work dW caused by a force \mathbf{F} acting for a distance $d\mathbf{r}$ is

$$dW = \mathbf{F} \cdot d\mathbf{r}$$

so the total work required to move a particle from point P_1 to point P_2 along a path γ is

$$W = \int_{\gamma} \mathbf{F} \cdot d\mathbf{r}.$$

The kinetic energy K is quantified by the quadratic form

$$K = \frac{1}{2}m|\mathbf{v}|^2 = \frac{1}{2}\mathbf{v} \cdot \mathbf{v}.$$

The change in potential energy is defined in terms of the work required to move the particle from P_1 to P_2 in a conservative force field — hence $\oint \mathbf{F} \cdot d\mathbf{r} = 0$ for any closed path. If we denote the potential energy at the endpoints by U_1 and U_2 , then

$$U_2 - U_1 = -W.$$

For conservative forces \mathbf{F} , the potential energy is related to the force by the gradient relation

$$\mathbf{F} = -\nabla U. \tag{C.7}$$

In 1-D, one can integrate (C.7) to obtain

$$U(x) = - \int_{x_0}^x F(s)ds;$$

however, in 2-D and 3-D this is not always possible. The negative sign in these relations can be motivated by the observation that if \mathbf{F} denotes the force due to gravity, the potential energy increases as the particle is lifted.

The *total energy* is the sum

$$\begin{aligned} \mathcal{H} &= K + U \\ &= \frac{1}{2}m\dot{\mathbf{r}} \cdot \dot{\mathbf{r}} + U \end{aligned}$$

of the kinetic and potential energies. For conservative forces in 3-D, it is observed that

$$\begin{aligned} \frac{\partial \mathcal{H}}{\partial t} &= m\ddot{\mathbf{r}} \cdot \dot{\mathbf{r}} + \sum_{i=1}^3 \frac{\partial U}{\partial r_i} \frac{dr_i}{dt} \\ &= (m\ddot{\mathbf{r}} - \mathbf{F}) \cdot \dot{\mathbf{r}} \\ &= 0. \end{aligned}$$

This expresses the *law of energy conservation* for conservative systems.

Lagrangian Mechanics

The development of Lagrangian theory for mechanical systems is based on the observation that variational principles form the basis for several of the fundamental laws obtained from Newtonian principles — e.g., force and moment balancing. To illustrate, consider the functional (C.3),

$$J = \int_{t_0}^{t_1} \mathcal{L}(\mathbf{r}, \dot{\mathbf{r}}, t)dt,$$

444 Appendix C. Legendre Transforms, Calculus of Variations, Mechanics Principles

where the Lagrangian

$$\mathcal{L} = K - U$$

is taken to be the difference between the kinetic and potential energies. It was shown in (C.5) that the extremum satisfies the Euler–Lagrange equations

$$\frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \dot{r}_i} - \frac{\partial \mathcal{L}}{\partial r_i} = 0$$

for $i = 1, \dots, 3$. Noting that $U = U(\mathbf{r})$ and $K = \frac{1}{2}m \sum_{i=1}^3 \dot{r}_i^2$, it follows that

$$\frac{\partial \mathcal{L}}{\partial r_i} = -\frac{\partial U}{\partial r_i}, \quad \frac{\partial \mathcal{L}}{\partial \dot{r}_i} = m\dot{r}_i$$

where the latter relation defines the momentum in terms of the Lagrangian. Hence the Euler–Lagrange equations yield

$$m\ddot{\mathbf{r}} = \mathbf{F}$$

which is precisely Newton’s second law.

Until now, we have employed rectangular coordinates when summarizing fundamental physical principles. For many systems, however, other coordinates may be more natural — e.g., polar or spherical. Hence it is common to employ a minimal number of *generalized coordinates*

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

required to specify the motion of a particle, body, or system. The following definition generalizes several of the concepts previously discussed in the context of a point mass in rectangular coordinates.

Definition C.3.1. *All relations hold for $i = 1, \dots, n$.*

$\mathbf{r} = \mathbf{r}(q_1, \dots, q_n, t)$	<i>Position vector for the body</i>
$\dot{q}_i = \frac{dq_i}{dt}$	<i>Generalized velocities</i>
$\mathcal{L}(\mathbf{q}, \dot{\mathbf{q}}, t) = K - U$	<i>Lagrangian</i>
$\mathcal{A} = \int_{t_0}^{t_1} \mathcal{L}(\mathbf{q}, \dot{\mathbf{q}}, t) dt$	<i>Action integral</i>
$\frac{\partial \mathcal{L}}{\partial q_i}$	<i>Generalized forces</i>
$\frac{\partial \mathcal{L}}{\partial \dot{q}_i} = p_i$	<i>Generalized or conjugate momenta</i>
$\frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \dot{q}_i} - \frac{\partial \mathcal{L}}{\partial q_i} = 0$	<i>Euler–Lagrange equations</i>

We note that in rectangular coordinates, the generalized momenta p_i are precisely the linear momenta $m\dot{x}_i$ whereas they are the angular momenta in polar coordinates. For arbitrary choices of generalized coordinates, the physical interpretation of p_i is less direct.

Details regarding the physics embodied in the Lagrangian framework can be found in [15] and additional theory is provided in [319].

Hamiltonian Mechanics

Whereas Lagrangian mechanics is based on a variational interpretation of physical principles, the Hamiltonian framework relies on total energy principles. In the former, the Lagrangian

$$\mathcal{L}(\mathbf{q}, \dot{\mathbf{q}}, t) = K(\dot{\mathbf{q}}) - U(\mathbf{q}, t),$$

defined in terms of generalized coordinates and their derivatives, provides the fundamental function used to quantify physical properties. In the Hamiltonian framework, the Hamiltonian

$$\mathcal{H}(\mathbf{q}, \mathbf{p}, t) = \dot{\mathbf{q}} \cdot \mathbf{p} - \mathcal{L}(\mathbf{q}, \dot{\mathbf{q}}, t), \tag{C.8}$$

where $p_i = \frac{\partial \mathcal{L}}{\partial \dot{q}_i}$ are conjugate or generalized momenta, is the fundamental quantity. From (C.2), it is observed that \mathcal{H} is the Legendre transform of \mathcal{L} .

The differential of (C.8) is

$$\begin{aligned} d\mathcal{H} &= \sum_{i=1}^n \left[p_i d\dot{q}_i + \dot{q}_i dp_i - \frac{\partial \mathcal{L}}{\partial q_i} dq_i - \frac{\partial \mathcal{L}}{\partial \dot{q}_i} d\dot{q}_i \right] - \frac{\partial \mathcal{L}}{\partial t} dt \\ &= \sum_{i=1}^n [p_i d\dot{q}_i + \dot{q}_i dp_i - \dot{p}_i dq_i - p_i d\dot{q}_i] - \frac{\partial \mathcal{L}}{\partial t} dt \\ &= \sum_{i=1}^n [\dot{q}_i dp_i - \dot{p}_i dq_i] - \frac{\partial \mathcal{L}}{\partial t} dt \end{aligned} \tag{C.9}$$

where the second step results from the definition $p_i = \frac{\partial \mathcal{L}}{\partial \dot{q}_i}$ and identity $\dot{p}_i = \frac{\partial \mathcal{L}}{\partial q_i}$ resulting from the Euler–Lagrange equations. Equating (C.9) with the total differential

$$d\mathcal{H} = \sum_{i=1}^n \left[\frac{\partial \mathcal{H}}{\partial q_i} dq_i + \frac{\partial \mathcal{H}}{\partial p_i} dp_i \right] + \frac{\partial \mathcal{H}}{\partial t} dt$$

yields Hamilton’s equations

$$\dot{q}_i = \frac{\partial \mathcal{H}}{\partial p_i} \quad , \quad \dot{p}_i = -\frac{\partial \mathcal{H}}{\partial q_i} \quad , \quad \frac{\partial \mathcal{H}}{\partial t} = -\frac{\partial \mathcal{L}}{\partial t} \quad , \quad i = 1, \dots, n.$$

Note that the first-order Hamilton’s equations are equivalent to the second-order Euler–Lagrange equations.

446 Appendix C. Legendre Transforms, Calculus of Variations, Mechanics Principles

Example 5. To illustrate properties of the Hamiltonian and Hamilton's equations of motion, we consider the 1-D motion of a mass m in response to a conservative force F . In this case

$$\mathcal{L}(x, \dot{x}, t) = \frac{1}{2}m\dot{x}^2 - U(x)$$

and $p = \frac{\partial \mathcal{L}}{\partial \dot{x}} = m\dot{x}$ so that

$$\begin{aligned}\mathcal{H}(x, p) &= m\dot{x}^2 - \left(\frac{1}{2}m\dot{x}^2 - U(x)\right) \\ &= K(\dot{x}) + U(x) \\ &= \frac{p^2}{2m} + U(x).\end{aligned}$$

Hence it is observed that \mathcal{H} is the sum of the kinetic and potential energies which is true in general for conservative systems. Furthermore, Hamilton's equations are

$$\begin{aligned}\frac{\partial \mathcal{H}}{\partial p} &= \frac{p}{m} = \dot{x} \\ \frac{\partial \mathcal{H}}{\partial x} &= \frac{\partial U}{\partial x} = -\dot{p}.\end{aligned}$$

The first expression simply relates the generalized momentum to the velocity and the second is Newton's second law (C.6).

Whereas the Lagrangian framework has the advantage of a variational basis, the Hamiltonian perspective is advantageous for certain applications of perturbation theory (celestial mechanics) and the characterization of complex systems arising in statistical mechanics and ergodic theory. It has also proven fundamental in the development of theories for optics and quantum mechanics. Further details regarding the physics and theory of Hamiltonian mechanics can be found in [15, 319].