Купить книгу Optimal Control

Modern Birkhäuser Classics

BIRKHÄUSER

Optimal Control

Richard Vinter

Chapter 1

Overview

Everything should be made as simple as possible, but not any simpler.

– Albert Einstein

1.1 Optimal Control

Optimal Control emerged as a distinct field of research in the 1950s, to address in a unified fashion optimization problems arising in scheduling and the control of engineering devices, beyond the reach of traditional analytical and computational techniques. Aerospace engineering is an important source of such problems, and the relevance of Optimal Control to the American and Russian space programs gave powerful initial impetus to the field. A simple example is:

The Maximum Orbit Transfer Problem. A rocket vehicle is in a circular orbit. What is the radius of the largest possible coplanar orbit to which it can be transferred over a fixed period of time? See Figure 1.1.

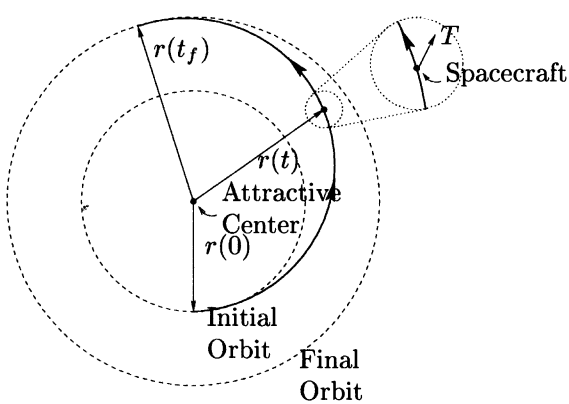

FIGURE 1.1. The Maximal Orbit Transfer Problem.

The motion of the vehicle during the maneuver is governed by the rocket thrust and by the rocket thrust orientation, both of which can vary with

R. Vinter, *Optimal Control*, Modern Birkhäuser Classics,
DOI 10.1007/978-0-8176-8086-2_1, © Springer Science+Business Media, LLC 2010

time. The variables involved are

$$\begin{array}{lcl} r & = & \text{radial distance of vehicle from attracting center,} \\ u & = & \text{radial component of velocity,} \\ v & = & \text{tangential component of velocity,} \\ m & = & \text{mass of vehicle,} \\ T_r & = & \text{radial component of thrust, and} \\ T_t & = & \text{tangential component of thrust.} \end{array}$$

The constants are

$$\begin{array}{lcl} r_0 & = & \text{initial radial distance,} \\ m_0 & = & \text{initial mass of vehicle,} \\ \gamma_{\max} & = & \text{maximum fuel consumption rate,} \\ T_{\max} & = & \text{maximum thrust,} \\ \mu & = & \text{gravitational constant of attracting center, and} \\ t_f & = & \text{duration of maneuver.} \end{array}$$

A precise formulation of the problem, based on an idealized point mass model of the space vehicle is as follows.

$$\left\{ \begin{array}{l} \text{Minimize } -r(t_f) \\ \text{over radial and tangential components of the thrust history,} \\ \qquad\qquad (T_r(t), T_t(t)),\ 0 \le t \le t_f, \text{ satisfying} \\ \dot r(t) = u, \\ \dot u(t) = v^2(t)/r(t) - \mu/r^2(t) + T_r(t)/m(t), \\ \dot v(t) = -u(t)v(t)/r(t) + T_t(t)/m(t), \\ \dot m(t) = -(\gamma_{\max}/T_{max})(T_r^2(t) + T_t^2(t))^{1/2}, \\ (T_r^2(t) + T_t^2(t))^{1/2} \le T_{\max}, \\ m(0) = m_0,\ r(0) = r_0,\ u(0) = 0,\ v(0) = \sqrt{\mu/r_0}, \\ u(t_f) = 0,\ v(t_f) = \sqrt{\mu/r(t_f)}. \end{array} \right.$$

Here $\dot r(t)$ denotes $dr(t)/dt$, etc. It is standard practice in Optimal Control to formulate optimization problems as minimization problems. Accordingly, the problem of maximizing the radius of the terminal orbit $r(t_f)$ is replaced by the equivalent problem of minimizing the "cost" $-r(t_f)$. Notice that knowledge of the *control function* or *strategy* $(T_r(t), T_t(t))$, $0 \le t \le t_f$ permits us to calculate the cost $-r(t_f)$: we solve the differential equations, for the specified boundary conditions at time $t = 0$, to obtain the corresponding *state trajectory* $(r(t), u(t), v(t), m(t))$, $0 \le t \le t_f$, and thence determine $-r(t_f)$. The control strategy therefore has the role of choice variable in the optimization problem. We seek a control strategy that minimizes the cost, from among the control strategies whose associated state trajectories satisfy the specified boundary conditions at time $t = t_f$.

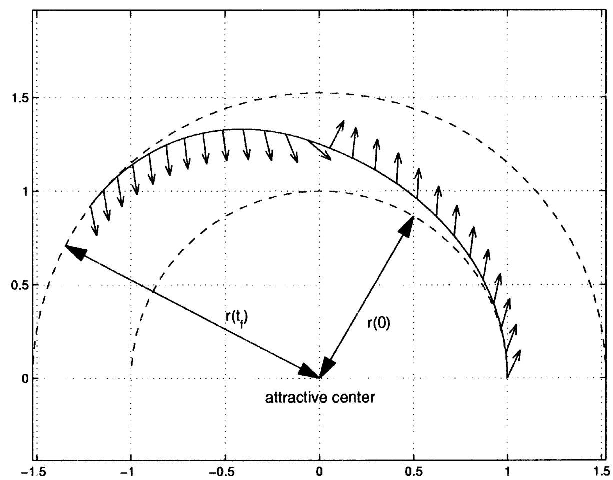

FIGURE 1.2. An Orbit Transfer Strategy.

Figure 1.2 shows a control strategy for the following values of relevant dimensionless parameters:

$$\frac{T_{\max}/m_0}{\mu/r_0^2} = 0.1405, \quad \frac{\gamma_{\max}}{T_{\max}/\sqrt{\mu/r_0}} = 0.07487, \quad \frac{t_f}{\sqrt{r_0^3/\mu}} = 3.32\,.$$

For this strategy, the radius of the terminal circular orbit is

$$r(t_f) = 1.5 r_0.$$

The arrows indicate the magnitude and orientation of the thrust at times $t = 0,\ 0.1t_f,\ 0.2t_f, \ldots,\ t_f$. As indicated, full thrust is maintained. The thrust is outward for (approximately) the first half of the maneuver and inward for the second.

Suppose, for example, that the attracting center is the Sun, the space vehicle weighs 10,000 lb, the initial radius is 1.50 million miles (the radius of a circle approximating the Earth's orbit), the maximum thrust is 0.85 lb (i.e., a force equivalent to the gravitational force on a 0.85 lb mass on the surface of the earth), the maximum rate of fuel consumption is 1.81 lb/day, and the transit time is 193 days. Corresponding values of the constants are

$$\begin{aligned} T_{\max} &= 3.778\,\mathrm{N}, & m_0 &= 4.536 \times 10^3\,\mathrm{kg},\\ r_0 &= 1.496 \times 10^{11}\,\mathrm{m}, & \gamma_{\max} &= 0.9496 \times 10^{-5}\,\mathrm{kg\,s^{-1}},\\ t_f &= 1.6675 \times 10^7\,\mathrm{s}, & \mu &= 1.32733 \times 10^{20}\,\mathrm{m^3 s^{-2}}. \end{aligned}$$

Then the terminal radius of the orbit is 2.44 million miles. (This is the radius of a circle approximating the orbit of the planet Mars.)

Numerical methods, inspired by necessary conditions of optimality akin to the Maximum Principle of Chapter 6, were used to generate the above control strategy.

Optimal Control has its origins in practical flight mechanics problems. But now, 40 years on, justification for research in this field rests not only on aerospace applications, but on applications in new areas such as process control, resource economics, and robotics. Equally significant is the stimulus Optimal Control has given to research in related branches of mathematics (convex analysis, nonlinear analysis, functional analysis, and dynamical systems).

From a modern perspective, Optimal Control is an outgrowth of the Calculus of Variations which takes account of new kinds of constraints (differential equation constraints, pathwise constraints on control functions "parameterizing" the differential equations, etc.) encountered in advanced engineering design. A number of key recent developments in Optimal Control have resulted from marrying old ideas from the Calculus of Variations and modern analytical techniques. For purposes both of setting Optimal Control in its historical context and of illuminating current developments, we pause to review relevant material from the classical Calculus of Variations.

1.2 The Calculus of Variations

The Basic Problem in the Calculus of Variations is that of finding an arc $\bar{x}$ which minimizes the value of an integral functional

$$J(x) = \int_S^T L(t, x(t), \dot{x}(t))dt$$

over some class of arcs satisfying the boundary condition

$$x(S) = x_0 \quad \text{and} \quad x(T) = x_1.$$

Here $[S,T]$ is a given interval, $L : [S,T] \times R^n \times R^n \to R$ is a given function, and x_0 and x_1 are given points in R^n.

The Brachistochrone Problem. An early example of such a problem was the *Brachistochrone Problem* circulated by Johann Bernoulli in the late 17th century. Positive numbers s_f and x_f are given. A frictionless bead, initially located at the point $(0,0)$, slides along a wire under the force of gravity. The wire, which is located in a fixed vertical plane, joins the points $(0,0)$ and (s_f, x_f). What should the shape of the wire be, in order that the bead arrive at its destination, the point (s_f, x_f), in minimum time?

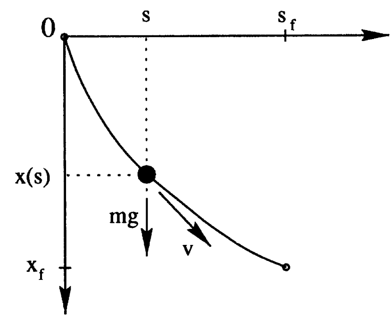

FIGURE 1.3. The Brachistochrone Problem.

There are a number of possible formulations of this problem. We now describe one of them. (See Figure 1.3.) Denote by s and x the horizontal and vertical distances of a point on the path of the bead (vertical distances are measured downward). We restrict attention to wires describable as the graph of a suitably regular function $x(s)$, $0 \leq s \leq s_f$. For any such function x, the velocity $v(s)$ is related to the downward displacement $x(s)$, when the horizontal displacement is s, according to

$$mgx(s) = \frac{1}{2}mv^2(s)$$

("loss of potential energy equals gain of kinetic energy"). But, denoting the time variable as t, we have

$$v(s) = \frac{\sqrt{1 + |dx(s)/ds|^2}}{dt(s)/ds} .$$

It follows that the transit time is

$$\int_0^{s_f} dt = \int_0^{s_f} \frac{\sqrt{1 + |dx(s)/ds|^2}}{v(s)} ds.$$

Eliminating $v(s)$ from the preceding expressions, we arrive at a formula for the transit time:

$$J(x) = \int_0^{s_f} L(s, x(s), \dot{x}(s))ds,$$

in which

$$L(s, x, w) := \frac{\sqrt{1 + |w|^2}}{\sqrt{2gx}}.$$

The problem is to minimize $J(x)$ over some class of arcs x satisfying

$$x(0) = 0 \text{ and } x(s_f) = x_f.$$

This is an example of the Basic Problem of the Calculus of Variations, in which $(S, x_0) = (0, 0)$ and $(T, x_1) = (s_f, x_f)$.

Suppose that we seek a minimizer in the class of absolutely continuous arcs. It can be shown that the minimum time t^* and the minimizing arc $(x(t), s(t))$, $0 \le t \le t^*$ (expressed in parametric form with independent variable time t) are given by the formulae

$$x(t) = a\left(1 - \cos\sqrt{\frac{g}{a}}\,t\right) \quad \text{and} \quad s(t) = a\left(\sqrt{\frac{g}{a}}\,t - \sin\sqrt{\frac{g}{a}}\,t\right).$$

Here, a and t^* are constants that uniquely satisfy the conditions

$$\begin{aligned} x(t^*) &= x_f, \\ s(t^*) &= t_f, \\ 0 \le & \sqrt{\frac{g}{a}}\,t^* \le 2\pi . \end{aligned}$$

The minimizing curve is a cycloid, with infinite slope at the point of departure: it coincides with the locus of a point on the circumference of a disc of radius a, which rolls without slipping along a line of length t_f.

Problems of this kind, the minimization of integral functionals, may perhaps have initially attracted attention as individual curiosities. But throughout the 18th and 19th centuries their significance became increasingly evident, as the list of the laws of physics that identified states of nature with minimizing curves and surfaces lengthened. Examples of "Rules of the Minimum" are as follows.

Fermat's Principle in Optics. The path of a light ray achieves a local minimum of the transit times over paths between specified endpoints that visit the relevant reflecting and refracting boundaries. The principle predicts Snell's Laws of Reflection and Refraction, and the curved paths of light rays in inhomogeneous media. See Figure 1.4.

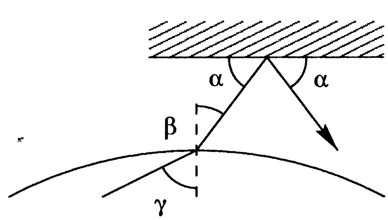

FIGURE 1.4. Fermat's Principle Predicts Snell's Laws.

Dirichlet's Principle. Take a bounded, open set $\Omega \subset R^2$ with boundary $\partial\Omega$, in which a static two-dimensional electric field is distributed. Denote by $V(x)$ the voltage at point $x \in \Omega$. Then $V(x)$ satisfies Poisson's Equation

$$\begin{aligned} \Delta V(x) &= 0 \text{ for } x \in \Omega \\ V(x) &= \bar{V}(x) \text{ for } x \in \partial\Omega. \end{aligned}$$

Here, $\bar{V} : \partial\Omega \to R$ is a given function, which supplies the boundary data.

Dirichlet's Principle characterizes the solution to this partial differential equation as the solution of a minimization problem

$$\begin{cases} \text{Minimize } \int_\Omega \nabla V(x) \cdot \nabla V(x) dx \\ \text{over surfaces } V \text{ satisfying } V(x) = \bar{V}(x) \text{ on } \partial\Omega. \end{cases}$$

This optimization problem involves finding a *surface* that minimizes a given integral functional. See Figure 1.5.

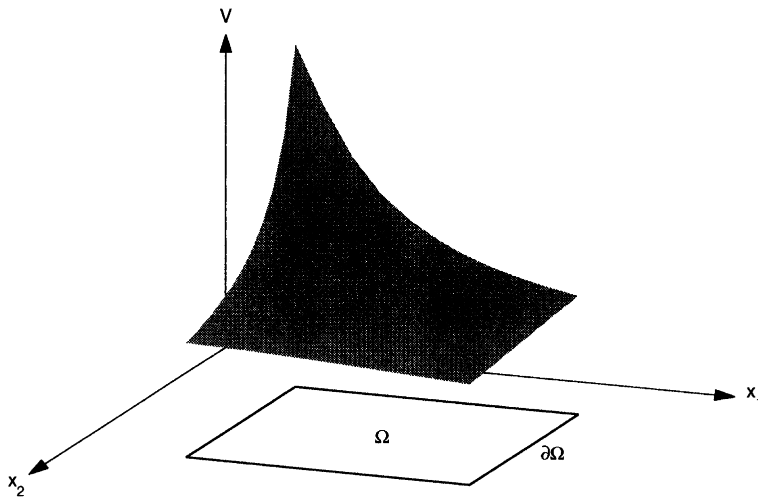

FIGURE 1.5. A Mimimizer for the Dirichlet Integral.

Dirichlet's Principle and its generalizations are important in many respects. They are powerful tools for the study of existence and regularity of solutions to boundary value problems. Furthermore, they point the way to Galerkin methods for computing solutions to partial differential equations, such as Poisson's equation: the solution is approximated by the minimizer of the above Dirichlet integral over some finite-dimensional subspace $\mathcal{S}_N$ of the domain of the original optimization problem, spanned by a finite collection of "basis" functions $\{\phi_i\}_{i=1}^N$,

$$\mathcal{S}_N = \{\sum_{i=1}^{N} \alpha_i \phi_i(x) : \alpha \in R^N\}.$$

The widely used finite element methods are modern implementations of Galerkin's method.

The Action Principle. Let $x(t)$ be the vector of generalized coordinates of a conservative mechanical system. The Action Principle asserts that $x(t)$

evolves in a manner to minimize (strictly speaking, to render stationary) the "action," namely,

$$\int [T(x(t), \dot{x}(t)) - V(x(t))]dt.$$

Here $T(x, \dot{x})$ is the kinetic energy and $V(x)$ is the potential energy. Suppose, for example, $x = (r, \theta)$, the polar coordinates of an object of mass m moving in a plane under the influence of a radial field (the origin is the center of gravity of a body, massive in relation to the object). See Figure 1.6. Then

$$T(x, \dot{x}) = \frac{1}{2}m(\dot{r}^2 + r^2\dot{\theta}^2)$$

and

$$V(r) = -K/r,$$

for some constant K. The action in this case is

$$\int \left(\frac{1}{2}m[\dot{r}^2(t) + r^2(t)\dot{\theta}^2(t)] - K/r(t)\right) dt.$$

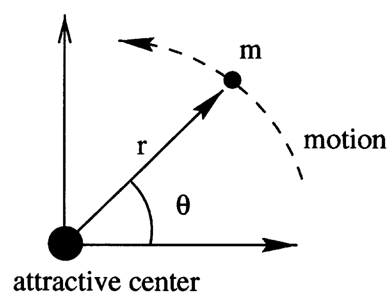

FIGURE 1.6. Motion in the Plane for a Radial Field.

The Action Principle has proved a fruitful starting point for deriving the dynamical equations of complex interacting systems, and for studying their qualitative properties (existence of periodic orbits with prescribed energy, etc.).

Necessary Conditions

Consider the optimization problem

$$(\text{CV}) \quad \begin{cases} \text{Minimize } J(x) := \int_S^T L(t, x(t), \dot{x}(t))dt \\ \text{over arcs } x \text{ satisfying} \\ x(S) = x_0 \text{ and } x(T) = x_1, \end{cases}$$

in which $[S,T]$ is a fixed interval, $L : R \times R^n \times R^n \to R$ is a given C^2 function, and x_0 and x_1 are given points in R^n. Precise formulation of problem (CV) requires us to specify the domain of the optimization problem. We take this to be $W^{1,1}([S,T];R^n)$, the class of absolutely continuous R^n-valued arcs on $[S,T]$, for reasons that are discussed presently.

The systematic study of minimizers $\bar{x}$ for this problem was initiated by Euler, whose seminal paper of 1744 provided the link with the equation:

$$\frac{d}{dt}L_v(t,\bar{x}(t),\dot{\bar{x}}(t)) = L_x(t,\bar{x}(t),\dot{\bar{x}}(t)). \tag{1.1}$$

(In this equation, L_x and L_v are the gradients of $L(t,x,v)$ with respect to the second and third arguments, respectively.)

The Euler Equation (1.1) is, under appropriate hypotheses, a necessary condition for an arc $\bar{x}$ to be a minimizer. Notice that, if the minimizer $\bar{x}$ is a C^2 function, then the Euler Equation is a second-order, n-vector differential equation:

$$\begin{aligned} L_{vt}(t,\bar{x}(t),\dot{\bar{x}}(t)) \;+\; L_{vx}(t,\bar{x}(t),\dot{\bar{x}}(t))\,\dot{\bar{x}}(t) \\ +\, L_{vv}(t,\bar{x}(t),\dot{\bar{x}}(t))\;\ddot{\bar{x}}\,(t) \;=\; L_x(t,\bar{x}(t),\dot{\bar{x}}(t)). \end{aligned}$$

A standard technique for deriving the Euler Equation is to reduce the problem to a scalar optimization problem, by considering a one-parameter family of *variations*. The Calculus of Variations, incidentally, owes its name to these ideas. (*Variations* of the minimizing arc cannot reduce the cost; conditions on minimizers are then derived by processing this information, with the help of a suitable *calculus* to derive necessary conditions of optimality.) Because of its historical importance and its continuing influence on the derivation of necessary conditions, we now describe the technique in detail.

Fix attention on a minimizer $\bar{x}$. Further hypotheses are required to derive the Euler Equation. We assume that there exists some number K such that

$$|L(t,x,v) - L(t,y,w)| \;\le\; K(|x-y| + |v-w|) \tag{1.2}$$

for all $x, y \in R^n$ and all $v, w \in R^n$.

Take an arbitrary C^1 arc y, which satisfies the homogeneous boundary conditions

$$y(S) = y(T) = 0.$$

Then, for any $\epsilon > 0$, the "variation" $x + \epsilon y$, which satisfies the endpoint constraints, must have cost not less than that of $\bar{x}$. It follows that

$$\epsilon^{-1}[J(\bar{x}+\epsilon y) - J(\bar{x})] \;\ge\; 0.$$

Otherwise expressed

$$\int_S^T \epsilon^{-1}[L(t,\bar{x}(t)+\epsilon y(t),\dot{\bar{x}}(t)+\epsilon\dot{y}(t)) - L(t,\bar{x}(t),\dot{\bar{x}}(t))]dt \;\ge\; 0.$$

Under Hypothesis (1.2), the Dominated Convergence Theorem permits us to pass to the limit under the integral sign. We thereby obtain the equation

$$\int_S^T [L_x(t, \bar{x}(t), \dot{\bar{x}}(t)) \cdot y(t) + L_v(t, \bar{x}(t), \dot{\bar{x}}(t)) \cdot \dot{y}(t)]dt \geq 0.$$

This relationship holds, we note, for all continuously differentiable functions y satisfying the boundary conditions $y(S) = 0$ and $y(T) = 0$. By homogeneity, the inequality can be replaced by equality.

Now apply integration by parts to the first term on the left. This gives

$$\int_S^T [-\int_S^t L_x(s, \bar{x}(s), \dot{\bar{x}}(s))ds + L_v(t, \bar{x}(t), \dot{\bar{x}}(t))] \cdot \dot{y}(t)dt = 0.$$

Take any continuous function $w : [S,T] \to R^n$ that satisfies

$$\int_S^T w(t)dt = 0. \tag{1.3}$$

Then the continuously differentiable arc $y(t) \equiv \int_S^t w(s)ds$ vanishes at the endtimes. Consequently,

$$\int_S^T [-\int_S^t L_x(s, \bar{x}(s), \dot{\bar{x}}(s))ds + L_v(t, \bar{x}(t), \dot{\bar{x}}(t))] \cdot w(t)dt = 0, \tag{1.4}$$

a relationship that holds for all continuous arcs w satisfying (1.3). To advance the analysis, we require

Lemma (Raymond Dubois) *Take a function $a \in L^2([S,T]; R^n)$. Suppose that*

$$\int_S^T a(t) \cdot w(t)\, dt = 0 \tag{1.5}$$

for every continuous function w that satisfies

$$\int_S^T w(t)dt = 0. \tag{1.6}$$

Then there exists some vector $d \in R^n$ such that

$$a(t) = d \quad \text{for a.e. } t \in [S,T].$$

Proof. We give a contemporary proof (based on an application of the Separation Principle in Hilbert space) of this classical lemma, a precursor of 20th century theorems on the representation of distributions.

By hypothesis, for any continuous function w, (1.6) implies (1.5). By using the fact that the continuous functions are dense in $L^2([S,T]; R^n)$

with respect to the strong L^2 topology, we readily deduce that (1.6) implies (1.5) also when w is a L^2 function. This fact is used shortly.

Define the constant elements $e^j \in L^2([S,T];R^n)$, $j = 1,\dots,n$ to have components:

$$e_i^j(t) = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}, \quad i = 1,\dots,n.$$

Since a function is a constant function if and only if it can be expressed as a linear combination of the $e_i^j(t)$s, the properties asserted in the lemma can be rephrased:

$$a \in V,$$

where V is the (closed) n-dimensional subspace of $L^2([S,T];R^n)$ spanned by the e^js. Suppose that the assertions are false; i.e., $a \notin V$. Since V is a closed subspace, we deduce from the Separation Theorem that there exists a nonzero element $w \in L^2$ and $\epsilon > 0$ such that

$$< a, w >_{L^2} \quad \leq \quad < v, w >_{L^2} - \epsilon.$$

for all $v \in V$. Because V is a subspace, it follows that

$$< v, w >_{L^2} \; = \; 0$$

for all $v \in V$. But then

$$< a, w >_{L^2} \leq \; -\epsilon. \tag{1.7}$$

Observe that, for each j,

$$< e^j, w >_{L^2} = 0,$$

since $e^j \in V$. This last condition can be expressed as

$$\int_S^T w(t)dt \; = \; 0.$$

In view of our earlier comments,

$$< a, w >_{L^2} = \; 0.$$

This contradicts (1.7). The lemma is proved. □

Return now to the derivation of the Euler Equation. We identify the function $a(.)$ of the lemma with

$$t \to -\int_S^t L_x(s, \bar{x}(s), \dot{\bar{x}}(s))ds + L_v(t, \bar{x}(t), \dot{\bar{x}}(t)).$$

In view of (1.4), the lemma informs us that there exists a vector d such that

$$-\int_S^t L_x(s,\bar{x}(s),\dot{\bar{x}}(s))ds + L_v(t,\bar{x}(t),\dot{\bar{x}}(t)) \;=\; d \text{ a.e.} \tag{1.8}$$

Since $L_x(t,\bar{x}(t),\dot{\bar{x}}(t))$ is integrable, it follows that $t \to L_v(t,\bar{x}(t),\dot{\bar{x}}(t))$ is almost everywhere equal to an absolutely continuous function and

$$\frac{d}{dt}L_v(t,\bar{x}(t),\dot{\bar{x}}(t)) = L_x(t,\bar{x}(t),\dot{\bar{x}}(t)) \text{ a.e.}$$

We have verified the Euler Equation and given it a precise interpretation, when the domain of the optimization problem is the class of absolutely continuous arcs.

The above analysis conflates arguments assembled over several centuries. The first step was to show that smooth minimizers $\bar{x}$ satisfy the pointwise Euler Equation. Euler's original derivation made use of discrete approximation techniques. Lagrange's alternative derivation introduced variational methods similar to those outlined above (though differing in the precise nature of the "integration by parts" step). Erdmann subsequently discovered that, if the domain of the optimization problem is taken to be the class of piecewise C^1 functions (i.e., absolutely continuous functions with piecewise continuous derivatives) then

"the function $t \to L_v(t,\bar{x}(t),\dot{\bar{x}}(t))$ has removable discontinuities."

This condition is referred to as the *First Erdmann Condition*.

For piecewise C^1 minimizers, the integral version of the Euler Equation (1.8) was first regarded as a convenient way of combining the pointwise Euler Equation and the First Erdmann Condition. We refer to it as the *Euler–Lagrange Condition*. An analysis in which absolutely continuous minimizers substitute for piecewise C^1 minimizers is an early 20th century development, due to Tonelli.

Another important property of minimizers can be derived in situations when L is independent of the t (write $L(x,v)$ in place of $L(t,x,v)$). In this "autonomous" case the second-order n-vector differential equation of Euler can be integrated. There results a first-order differential equation involving a "constant of integration" c:

$$L_v(\bar{x}(t),\dot{\bar{x}}(t)) \cdot \dot{\bar{x}}(t) - L(\bar{x}(t),\dot{\bar{x}}(t)) = c\,. \tag{1.9}$$

This condition is referred to as the *Second Erdmann Condition* or *Constancy of the Hamiltonian Condition.* It is easily deduced from the Euler–Lagrange Condition when $\bar{x}$ is a C^2 function. Fix t. We calculate in this case

$$\frac{d}{dt}(L_v \cdot \dot{\bar{x}}(t) - L) \;=\; \frac{d}{dt}L_v \cdot \dot{\bar{x}}(t) + L_v\cdot \ddot{\bar{x}}\,(t) - L_x \cdot \dot{\bar{x}}(t) - L_v\cdot \ddot{\bar{x}}\,(t)$$

$$
\begin{aligned}
&= (\frac{d}{dt}L_v - L_x) \cdot \dot{\bar{x}}(t) \\
&= 0.
\end{aligned}
$$

(In the above relationships, L, L_v, etc., are evaluated at $(\bar{x}(t), \dot{\bar{x}}(t))$.) We deduce (1.9).

A more sophisticated analysis leads to an "almost everywhere"' version of this condition for autonomous problems, when the minimizer $\bar{x}$ in question is assumed to be merely absolutely continuous.

Variations of the type $\bar{x}(t) + \epsilon y(t)$ lead to the Euler–Lagrange Condition. Necessary conditions supplying additional information about minimizers have been derived by considering other kinds of variations. We note in particular the *Weierstrass Condition* or the *Constancy of the Hamiltonian Condition*:

$$
\begin{aligned}
&L_v(t, \bar{x}(t), \dot{\bar{x}}(t)) \cdot \dot{\bar{x}}(t) - L(t, \bar{x}(t), \dot{\bar{x}}(t)) \\
&\quad = \max_{v \in R^n} \{L_v(t, \bar{x}(t), \dot{\bar{x}}(t)) \cdot v - L(t, \bar{x}(t), v)\}.
\end{aligned}
$$

Suppressing the $(t, \bar{x}(t))$ argument in the notation and expressing the Weierstrass Condition as

$$
L(v) - L(\dot{\bar{x}}(t)) \geq L_v(\dot{\bar{x}}) \cdot (v - \dot{\bar{x}}(t)) \quad \text{for all } v \in R^n,
$$

we see that it conveys no useful information when $L(t, x, v)$ is convex with respect to the v variable: in this case it simply interprets L_v as a subgradient of L in the sense of convex analysis. In general however, it tells us that L coincides with its "convexification" (with respect to the velocity variable) along the optimal trajectory; i.e.,

$$
L(t, \bar{x}(t), \dot{\bar{x}}(t)) = \tilde{L}(t, \bar{x}(t), \dot{\bar{x}}(t)).
$$

Here $\tilde{L}(t, x, .)$ is the function with epigraph set co $\{\text{epi}\, L(t, x, .)\}$. See Figure 1.7.

The above necessary conditions (the Euler–Lagrange Condition, the Weierstrass Condition, and the Second Erdmann Condition) have convenient formulations in terms of the "adjoint arc"

$$
p(t) = L_v(t, \bar{x}(t), \dot{\bar{x}}(t)).
$$

They are

$$
(\dot{p}(t), p(t)) = \nabla_{x,v} L(t, \bar{x}(t), \dot{\bar{x}}(t)), \tag{1.10}
$$

$$
p(t) \cdot \dot{\bar{x}}(t) - L(t, \bar{x}(t), \dot{\bar{x}}(t)) = \max_{v \in R^n} [p(t) \cdot v - L(t, \bar{x}(t), v)],
$$

and, in the case when L does not depend on t,

$$
p(t) \cdot \dot{\bar{x}}(t) - L(t, \bar{x}(t), \dot{\bar{x}}(t)) = c
$$

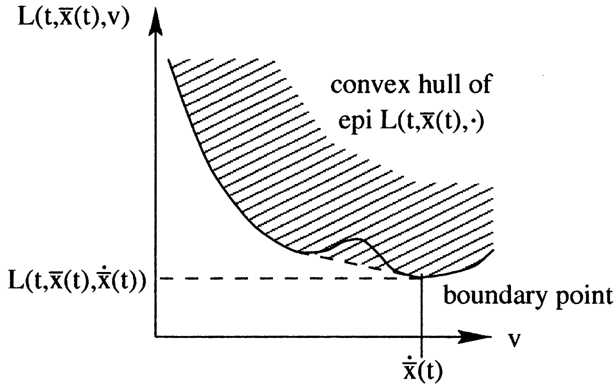

FIGURE 1.7. The Weierstrass Condition.

for some constant c.

To explore the qualitative properties of minimizers it is often helpful to reduce the Euler–Lagrange Condition to a system of specially structured first-order differential equations. This was first achieved by Hamilton for Lagrangians L arising in mechanics. In this analysis the *Hamiltonian*,

$$H(t,x,p) := \max_{v \in R^n} \{p \cdot v - L(t,x,v)\},$$

has an important role.

Suppose that the right side has a unique maximizer v_{max}. This will depend on (t,x,p) and so we can use it to define a function $\chi : [S,T] \times R^n \times R^n \to R^n$:

$$\chi(t,x,p) := v_{\max}.$$

Then

$$H(t,x,p) = p \cdot \chi(t,x,p) - L(t,x,\chi(t,x,p))$$

and

$$\nabla_v (p \cdot v - L(t,x,v))|_{v=\chi(t,x,p)} = 0.$$

The last relationship implies

$$p = L_v(t,x,\chi(t,x,p)). \tag{1.11}$$

(The mapping from x vectors to p vectors implicit in this relationship, for fixed t, is referred to as the *Legendre Transformation.*)

Now consider an arc $\bar{x}$ and associated adjoint arc p that satisfy the Euler–Lagrange and Weierstrass Conditions, namely

$$(\dot{p}(t), p(t)) = \nabla_{x,v} L(t, \bar{x}(t), \dot{\bar{x}}(t)) \tag{1.12}$$

and

$$p(t) \cdot \dot{\bar{x}}(t) - L(t,\bar{x}(t),\dot{\bar{x}}(t)) = \max{}_{v \in R^n} \{p(t) \cdot v - L(t,\bar{x}(t),v)\}.$$

Since it is assumed that the "maximizer" in the definition of the Hamiltonian is unique, it follows from the Weierstrass condition that

$$\dot{\bar{x}}(t) = \chi(t, \bar{x}(t), p(t)).$$

Fix t. Let us assume that $\chi(t, ., .)$ is differentiable. Then we can calculate the gradients of $H(t, ., .)$:

$$\begin{aligned}
&\nabla_x H(t,x,p)|_{x=\bar{x}(t),p=p(t)} \\
&\quad = \nabla_x (p \cdot \chi(t,x,p) - L(t,x,\chi(t,x,p)))|_{x=\bar{x}(t),p=p(t)} \\
&\quad = p \cdot \chi_x(t,x,p) - L_x(t,x,\chi(t,x,p)) \\
&\qquad\qquad - L_v(t,x,\chi(t,x,p)) \cdot \chi_x(t,x,p)|_{x=\bar{x}(t),p=p(t)} \\
&\quad = (p - L_v(t,x,\chi(t,x,p))) \cdot \chi_x(t,x,p) - L_x(t,x,\chi(t,x,p))|_{x=\bar{x}(t),p=p(t)} \\
&\quad = 0 - \dot{p}(t).
\end{aligned}$$

(The last step in the derivation of these relationships makes use of (1.11) and (1.12).) We have evaluated the x-derivative of H:

$$\nabla_x H(t, \bar{x}(t), p(t)) \;=\; -\dot{p}(t).$$

As for the p-derivative, we have

$$\begin{aligned}
&\nabla_p H(t,x,p)|_{x=\bar{x}(t),p=p(t)} \\
&\quad = \nabla_p (p \cdot \chi(t,x,p) - L(t,x,\chi(t,x,p)))|_{x=\bar{x}(t),p=p(t)} \\
&\quad = \chi(t,x,p) + p \cdot \chi_p(t,x,p) - L_v(t,x,\chi(t,x,p)) \cdot \chi_p(t,x,p)|_{x=\bar{x}(t),p=p(t)} \\
&\quad = \dot{\bar{x}}(t) + (p(t) - L_v(t,\bar{x}(t),\dot{\bar{x}}(t))) \cdot \chi_p(t,\bar{x}(t),p(t)) \\
&\quad = \dot{\bar{x}}(t) + 0.
\end{aligned}$$

Combining these relationships, we arrive at the system of first-order differential equations of interest, namely, the *Hamilton Condition*

$$(-\dot{p}(t), \dot{\bar{x}}(t)) = \nabla H(t, \bar{x}(t), p(t)), \tag{1.13}$$

in which ∇H denotes the gradient of $H(t, ., .)$.

So far, we have limited attention to the problem of minimizing an integral functional of arcs with fixed endpoints. A more general problem is that in which the arcs are constrained to satisfy the boundary condition

$$(x(S), x(T)) \in C, \tag{1.14}$$

for some specified subset $C \subset R^n \times R^n$. (The fixed endpoint problem is a special case, in which C is chosen to be $\{(x_0, x_1)\}$. The above necessary conditions remain valid when we pass to this more general endpoint constraint, since a minimizer $\bar{x}$ over arcs satisfying (1.14) is also a minimizer for the fixed endpoint problem in which $C = \{\bar{x}(S), \bar{x}(T)\}$. But something