Lecture Notes on Dynamic Optimization

Piermarco Cannarsa & Elena Giorgieri & M. Elisabetta Tessitore

February 26, 2009

Contents

Notations

$C^k([a,b];\mathbb{R}^n)$	space of k times continuously differentiable functions $f:[a,b]\to\mathbb{R}^n$
$C^k([a,b])$	space of k times continuously differentiable functions $f:[a,b]\to\mathbb{R}$
$C_*([a,b];\mathbb{R}^n)$	space of piecewise continuous functions $f:[a,b]\to\mathbb{R}^n$
$C_*^1([a,b];\mathbb{R}^n)$	space of piecewise continuously differentiable functions $f:[a,b]\to\mathbb{R}^n$
$\frac{\partial f}{\partial x};\ D_x f;\ f_x$	partial derivatives of f with respect to the x variable
$\frac{\partial^2 f}{\partial x^2};\ D_x^2 f;\ f_{xx}$	second partial derivatives of f with respect to the x variable
$< a,b >;\ a\cdot b$	scalar product of a and b
$M_{n\times m}$	space of $n\times m$ matrices

Preface

The origin of Dynamic Optimization as a mathematical discipline can be traced back at least to the year 1696, when the first official problem in Calculus of Variations was formulated in a celebrated work by J. Bernouilli on the brachistocrone problem. However, such a subject must have interested mankind for a much longer time, at least according to Euler who wrote that 'nothing at all takes place in the universe in which some rule of maximum or minimum does not appear'. Since then, Calculus of Variations has been an extremely active research area, rich of surprising results and new stimulating problems, well connected with science and engineering. It was mostly for the needs of airspace engineering and economics that, with the rapid development of Optimal Control and Game Theory, in the second half of the twentieth century the subject grew into what is now considered to be Dynamic Optimization. Several adaptations of the theory were later required, including extensions to stochastic models and infinite dimensional processes.

These lecture notes are intended as a friendly introduction to Calculus of Variations and Optimal Control, for students in science, engineering and economics with a general background in ordinary differential equations and calculus for functions of several real variables. In order to keep the exposition at a beginner's level, we have deliberately omitted any attempt to study the existence of optimal controls. In this way, no knowledge of Lebesgue's integration theory is required. Here, the main focus will be on optimality conditions—either necessary or sufficient—and examples. In fact, we will examine in detail both mathematical and economical models, such as the problem of Queen Dido, the time-optimal capture of a wandering particle, the Evans model for a monopolistic firm production and the Nordhaus model for optimal strategies in a democracy.

This book is organized as follows. Chapter 1 is concerned with Calculus of Variations. After introducing some classical examples, we quickly derive first order necessary conditions, including the Euler-Lagrange equations and Erdmann's condition. Then, we apply these results to the examples introduced before. Afterwards, we discuss second order optimality conditions and deduce Jacobi's theory. Finally, we analyse a particular case of constrained problem, namely the isoperimetric problem. In Chapter 2, we study optimal control problems. First, we obtain necessary optimality conditions in the form of Pontryagin's Maximum Principle for the Mayer problem. Then, we extend this result to the Bolza problem and to problems with terminal constraints. The Maximum Principle is used to study several different examples. Finally, after having introduced the basic objects of Dynamic Programming, namely the value function, the optimality principle and the Hamilton-Jacobi-Bellman equation, we show how to use this technique to construct optimal trajectories. To assist the reader who may here be confronted with partial differential equations for the first time, we have provided a short exposition of the classical method of characteristics. We conclude with an Appendix on the classical Legendre-Fenchel transform.

The current version of these notes has profited by remarks and contributions coming from students in mathematics and economics who were trained in this subject over the last few years at the University of Rome Tor Vergata. Moreover, in writing Chapter 1, we have been influenced by the lectures given by Francis Clarke this year at Tor Vergata. We will also be grateful to anyone who will spot misprints and errors in the monograph, reporting them to one of the e-mail addresses below.

Piermarco Cannarsa
cannarsa@mat.uniroma2.it

Elena Giorgieri
giorgier@mat.uniroma2.it

Maria Elisabetta Tessitore
tessitore@sefemeq.uniroma2.it

Chapter 1

Calculus of Variations

1.1 Basic Problem in Calculus of Variations

Calculus of Variations is a branch of mathematics starting between the end of the seventeenth century and the middle of the eighteenth. Some of the basic problems arisen at that time were the following.

Example 1.1.1 Given two points (a, A) and (b, B) $(a \neq b)$ in the plane $\mathbb{R}^2$, we want to find the regular curve joining them that has minimal length. We can represent any such a curve γ by means of a regular function

$$x(\cdot) : [a, b] \to \mathbb{R}$$

such that $x(a) = A$, $x(b) = B$, for which the length is given by the formula

$$\text{Length}(\gamma) = \int_a^b \sqrt{1 + x'(t)^2}\, dt =: J(x),$$

where $x'(\cdot)$ denotes the derivative of $x(\cdot)$.
It is easy to see that the problem concerns the minimization of the functional J over the aforementioned class of functions.

Remark 1.1.2 The problem is of course trivial in the plane, because (as we will see later) the unique solution is given by the segment joining (a, A) and (b, B); but this example becomes more meaningful (and more involved) if, for instance, we require the curve to be in the space $\mathbb{R}^3$ and to join points on a given surface (the problem of geodesic).

Example 1.1.3 (The soap bubble problem) Consider in the space $\mathbb{R}^3$ the circles

$$\begin{cases} y^2 + z^2 = A^2 \\ x = a. \end{cases}, \qquad \begin{cases} y^2 + z^2 = B^2 \\ x = b. \end{cases},$$

where $a \neq b$. Consider any regular curve in the xz-plane $\xi : [a, b] \to \mathbb{R}^3$, $\xi(x) = (x, 0, \alpha(x))$ such that $\alpha(a) = A$ and $\alpha(b) = B$ and the surface of revolution generated by ξ.

We want to minimize the area of the resulting surface among all the regular functions ξ defined above. But the area of any such a surface S is given by

$$\text{Area}(S) = 2\pi \int_a^b \alpha(x)\sqrt{1+\alpha'(x)^2}\,dx =: J(\alpha),$$

so that the problem deals with the minimization of the functional J over the class of regular functions α such that $\alpha(a) = A$ and $\alpha(b) = B$.

Example 1.1.4 Consider a particle $x(t)$ moving from time t_1 to time t_2 between two points A and B and subject to a conservative force $F(x(t)) = -\nabla V(x(t))$. Among all the (admissible) trajectories, we want to find the one that minimizes the "action", i.e. the functional

$$J(x) = \int_{t_1}^{t_2} \left[\frac{1}{2}m|x'(t)|^2 - V(x(t))\right] dt,$$

where m is the mass of the particle and $\frac{1}{2}m|x'(t)|^2$ is its kinetic energy. Roughly speaking, we want to find the trajectory that goes from A to B in time $t_2 - t_1$ with "minimal dissipation of energy".

Example 1.1.5 (The brachistochrone problem - G. Bernoulli, 1696) Consider a vertical plane π and two fixed points (a, A) and (b, B), with $a < b$ and $A > B$. Take any regular curve γ connecting (a, A) and (b, B) and a point P bound to this curve that slips from (a, A) with initial velocity $v_0 > 0$ to (b, B) subject to the gravity force only. Among all the curves γ we want to find the one along which the point P reaches (b, B) in minimal time. If we represent γ as a graph of a regular function $x \mapsto y(x)$ such that $y(a) = A$ and $y(b) = B$, it can be shown that the time for P to cover the distance (along γ) between (a, A) and (b, B) is given by

$$\text{Time}(\gamma) = \frac{1}{\sqrt{2g}} \int_a^b \frac{1}{\sqrt{H - y(x)}} \sqrt{1 + y'(x)^2}dx =: J(y),$$

where $H = A + v_0^2/2g$ and g is the gravity acceleration. Hence, the given problem can be formulated as the minimization problem of the functional J over the class of regular functions $y(\cdot)$ such that $y(a) = A$ and $y(b) = B$.

The previous examples show that a natural approach to solve many problems in Calculus of Variations is to formulate them as a problem of minimization and (provided we know that solutions exist) to look for *necessary conditions* which have to be satisfied by minimizers. Although the setup of these examples is the space of C^1 functions, we will present a more general case, that is the one defined on the space of piecewise C^1 functions. Let us start by defining the general setting.

Definition 1.1.6 *A continuous function $\xi : [a, b] \to \mathbb{R}^n$ is said to be piecewise C^1 if there exists a finite partition of $[a, b]$, say*

$$\Pi = \{t_0, t_1, \dots, t_N\}, \qquad a = t_0 < t_1 < \dots < t_N = b,$$

such that $\xi \in C^1([t_i, t_{i+1}], \mathbb{R}^n)$ for any $i = 0, \dots, N-1$, meaning that $\xi \in C^1((t_i, t_{i+1}), \mathbb{R}^n)$ and the right and left derivatives exist at t_i and t_{i+1}, respectively. In the sequel we will denote by $C^1_*([a,b], \mathbb{R}^n)$ the class of all piecewise C^1 functions.

Fix $a, b \in \mathbb{R}$, $a < b$, and let $S_a, S_b \subset \mathbb{R}^n$ be closed nonempty sets. Denote by $C^1_*([a,b]; \mathbb{R}^n)$ the class of all piecewise continuously differentiable arcs $\xi : [a,b] \to \mathbb{R}^n$. We define the set of admissible arcs by

$$A = \{\xi \in C^1_*([a,b]; \mathbb{R}^n) : \xi(a) \in S_a, \xi(b) \in S_b\}.$$

Moreover we define the functional J over A as

$$J(\xi) = \int_a^b L(t, \xi(t), \xi'(t))dt + \phi_a(\xi(a)) + \phi_b(\xi(b)), \tag{1.1}$$

where $L : [a,b] \times \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$, $\phi_a : \mathbb{R}^n \to \mathbb{R}$, $\phi_b : \mathbb{R}^n \to \mathbb{R}$ are given continuous functions called *running cost* (or *Lagrangian*), *initial cost* and *terminal cost*, respectively. Notice that the integral in (1.1) is well defined on A since the number of discontinuity points of the derivative of any element $\xi \in A$ is finite.

The general problem we are dealing with is the minimization problem of the functional J over the class of functions A, i.e.

$$\min \{J(\xi) \mid \xi \in A\}. \tag{1.2}$$

Definition 1.1.7 *An admissible arc $\xi^* \in C^1_*([a,b]; \mathbb{R}^n)$ is a* minimizer *(or a* solution*) of problem (1.2) if*

$$\xi^* \in A \qquad \textit{and} \qquad J(\xi^*) = \min \{J(\xi) : \xi \in A\}. \tag{1.3}$$

In the sequel it will be also useful to consider the notion of local minimizers.

Definition 1.1.8 *An admissible arc $\bar{\xi} \in C^1_*([a,b]; \mathbb{R}^n)$ is a* local minimizer *of problem (1.2) if there exists $\varepsilon > 0$ such that $J(\bar{\xi}) \le J(\xi)$ for any admissible $\xi \in C^1_*([a,b]; \mathbb{R}^n)$ satisfying*

$$\max_{t \in [a,b]} |\bar{\xi}(t) - \xi(t)| < \varepsilon \qquad \textit{and} \qquad \max_{i=1,\dots,N} \max_{t \in [t_i, t_{i+1}]} |\bar{\xi}'(t) - \xi'(t)| < \varepsilon,$$

where $\Pi = \{t_0, t_1, \dots, t_N\}$ is a common partition for $\bar{\xi}$ and ξ.

To simplify the notations, in the sequel we will set

$$\|\bar{\xi} - \xi\|_\infty := \max_{t \in [a,b]} |\bar{\xi}(t) - \xi(t)|$$

and

$$\|\bar{\xi}' - \xi'\|^*_\infty := \max_{i=1,\dots,N} \max_{t \in [t_i, t_{i+1}]} |\bar{\xi}'(t) - \xi'(t)|.$$

Observe that any minimizer of problem (1.2) is a local minimizer, while the converse is not true in general.

Remark 1.1.9 In the sequel we will be concerned with minimization problems only since a maximization problem can be rewritten as a minimization one recalling that given a function f, then

$$\max f = -\min(-f).$$

In Calculus of Variations, the first problem we face is to ensure the existence of a minimizer ξ^* of functional (1.1). From now on L, ϕ_a and ϕ_b will be required to satisfy at least the following assumptions

$$\begin{aligned} &\phi_a, \phi_b \in C^1(\mathbb{R}^n), \text{ and } L \in C^1([a,b] \times \mathbb{R}^n \times \mathbb{R}^n); \\ &\phi_a, \phi_b, L \geq -c, \text{ for some } c \in \mathbb{R}^+. \end{aligned} \tag{1.4}$$

In the first half of the 20th century Tonelli proved the existence of minimizers for problem (1.2) under the following (stronger) assumptions on L, ϕ_a and ϕ_b:

(E1) $\phi_a, \phi_b \in C^1(\mathbb{R}^n)$, and $L \in C^1([a,b] \times \mathbb{R}^n \times \mathbb{R}^n)$;

$$\begin{aligned} &\phi_a, \phi_b \geq -c \text{ and } L(t,x,q) \geq \theta(|q|) - c, \text{ for some } c \in \mathbb{R}^+ \\ &\text{and some nonnegative function } \theta \text{ such that } \lim_{r \to +\infty} \frac{\theta(r)}{r} = +\infty. \end{aligned} \tag{1.5}$$

(E2) L is twice differentiable with respect to q and $\dfrac{\partial^2 L}{\partial q^2} > 0$.

Remark 1.1.10 In assumption (E1) ϕ_a, ϕ_b are required to be bounded from below, while L has to be superlinear, that is it has a fast growth. Function θ is often referred to as the Nagumo function for functional J.
For instance assumption (E1) holds if there exist $p > 1$, $\nu > 0$ and $\lambda \in \mathbb{R}$ such that

$$L(t,x,q) \geq \nu |q|^p + \lambda.$$

In assumption (E2), L is required to be strictly convex with respect to the third variable q. If $q \in \mathbb{R}^n$, then condition $\frac{\partial^2 L}{\partial q^2} > 0$ means that the Hessian matrix is positive definite. Hypothesis (E2) can be strongly relaxed requiring L to be a convex function with respect to the third variable.

Theorem 1.1.11 (Tonelli's Theorem) *Assume (E1) and (E2). Moreover, if S_a or S_b is a bounded set, then problem (1.2) admits at least one solution.*

For the proof see [3].

Let us show that the theorem does not hold true (in general) if assumptions (E1) or (E2) are not satisfied.

Example 1.1.12 Assumption (E1) does not hold.
Let $L : [0,1] \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ be defined as $L(t,x,q) = \sqrt{x^2+q^2}$, $S_0 = \{0\}$, $S_1 = \{1\}$ and $\phi_0 = \phi_1 = 0$. (E1) is not satisfied since L is sublinear.
The problem is to minimize

$$\left\{ \int_0^1 \sqrt{\xi(t)^2 + \xi'(t)^2} dt \mid \xi \in A \right\},$$

where $A = \{\xi \in C^1_*([0,1];\mathbb{R}) : \xi(0) = 0, \xi(1) = 1\}$.
It is easy to see that

$$\inf \left\{ \int_0^1 \sqrt{\xi(t)^2 + \xi'(t)^2} dt \mid \xi \in A \right\} = 1,$$

since for all $\xi \in A$

$$\int_0^1 L(t,\xi(t),\xi'(t))dt = \int_0^1 \sqrt{\xi(t)^2 + \xi'(t)^2} dt \geq \int_0^1 |\xi'(t)| dt \geq \xi(1) - \xi(0) = 1.$$

Hence if there exists ξ^* where the minimum is attained, then $J[\xi^*] \geq 1$.
On the other hand, evaluating the functional on the maps $\xi_k(t) = t^k$, $k \in \mathbb{N}$, we have

$$\begin{aligned} \int_0^1 L(t,\xi_k(t),\xi_k'(t))dt &= \int_0^1 t^{k-1}\sqrt{t^2+k^2}dt \leq \sqrt{k^2+1}\left[\frac{t^k}{k}\right]_0^1 \\ &= \frac{\sqrt{k^2+1}}{k} \to 1 \quad \text{as} \quad k \to \infty. \end{aligned}$$

Hence $J[\xi^*] = 1$, which yields $\xi^* = 0$. The conclusion that the minimum cannot exist follows since the boundary conditions are not satisfied.

Example 1.1.13 Assumption (E2) does not hold.
Consider the map $L : [0,1] \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$, defined by $L(t,x,q) = |x|^2 + (|q|^2 - 1)^2$, $S_0 = \{0\}$, $S_1 = \{0\}$ and $\phi_0, \phi_1 \equiv 0$.
It is easy to see that $\inf\{J(\xi) : \xi \in A\} \geq 0$. Moreover, for the maps $\xi_k \in A$, $k \geq 2$, given by

$$\xi_k(t) = \begin{cases} t - \dfrac{j}{k} & \text{for } t \in \left[\dfrac{j}{k}, \dfrac{2j+1}{2k}\right], \ j = 0, \dots, k-1, \\ \\ -t + \dfrac{j+1}{k} & \text{for } t \in \left]\dfrac{2j+1}{2k}, \dfrac{j+1}{k}\right], \ j = 0, \dots, k-1, \end{cases}$$

we have $\int_0^1 L(s,\xi_k(s),\xi_k'(s))ds = \frac{1}{12k^2} \to 0$, as $k \to \infty$. Hence $\inf\{J(\xi) : \xi \in A\} = 0$. Hence, if a minimizer would exist, it should be the null function $\xi^* \equiv 0$, as a simple computation shows. But this cannot happen, since $J(\xi^*) = 1$.

Remark 1.1.14 It is important to stress that the existence result stated above, which is meant for the problem over the class of piecewise C^1 functions, holds true also for the "restricted" problem on the class of C^1 functions, without further assumptions. However, the introductory Examples 1.1.1-1.1.3-1.1.5 do not satisfy hypotheses (E1) and (E2), so that Tonelli's Theorem does not apply. Of course there are other theorems that guarantee existence of minimizers for those problems, but their formulation is much more involved than Tonelli's one and their analysis is far beyond the scope of this course. So whenever we handle again Examples 1.1.1-1.1.3-1.1.5 we will assume existence.

In the sequel, we will not consider problem (1.2) in the general form, but we will mainly analyse two subcases:

$$\min\left\{\int_a^b L(t,\xi(t),\xi'(t))dt \mid \xi \in C^1_*([a,b];\mathbb{R}^n), \xi(a)=\xi_a, \xi(b)=\xi_b\right\},$$

which is a problem with given initial and terminal points, and null costs;

$$\min\left\{\int_a^b L(t,\xi(t),\xi'(t))dt + \phi_a(\xi(a)) \mid \xi \in C^1_*([a,b];\mathbb{R}^n), \xi(b)=\xi_b\right\},$$

which is a problem with a free initial point, a fixed terminal point and a not identically null initial cost.

1.2 First Order Necessary Conditions

1.2.1 Euler equation

The standard technique applied in Calculus of Variations consists, first of all, in imposing some necessary conditions in order to select from the admissible set of arcs A a suitable subset of candidate minimizers, trying then to pick out a particular arc at which the minimum value of the functional J is attained.

For the reader's convenience, we begin by considering the simpler problem of minimization over the class of C^1 functions, i.e.

$$\min\left\{\int_a^b L(t,\xi(t),\xi'(t))dt \mid \xi \in C^1([a,b];\mathbb{R}^n), \xi(a)=\xi_a, \xi(b)=\xi_b\right\}, \tag{1.6}$$

which is a problem with fixed boundary conditions and null costs.
Analogously to the C^1_* case, we can define the notion of *minimizer* and *local minimizer* in the C^1 case.

Definition 1.2.1 *An admissible arc $\xi^* \in C^1([a,b];\mathbb{R}^n)$ is a* minimizer *(or a* solution*) of problem (1.6) if*

$$\xi^* \in A \qquad \text{and} \qquad J(\xi^*) = \min\{J(\xi) : \xi \in A\}, \tag{1.7}$$

where $A = \{\xi \in C^1([a,b];\mathbb{R}^n) : \xi(a)=\xi_a, \xi(b)=\xi_b\}$.
An admissible arc $\bar\xi \in C^1([a,b];\mathbb{R}^n)$ is a local minimizer *of problem (1.6) if there exists $\varepsilon>0$ such that $J(\bar\xi) \le J(\xi)$ for any admissible ξ satisfying*

$$\max_{t\in[a,b]} |\bar\xi(t)-\xi(t)| < \varepsilon \qquad \text{and} \qquad \max_{t\in[a,b]} |\bar\xi'(t)-\xi'(t)| < \varepsilon.$$

Let us first prove a technical lemma which will be useful in the sequel.
We denote by $\cdot$ or $< \cdot, \cdot >$ the scalar product in $\mathbb{R}^n$.

Lemma 1.2.2 *Let $f, g \in C([a,b];\mathbb{R}^n)$ satisfy*

$$\int_a^b [f(t) \cdot \eta(t) + g(t) \cdot \eta'(t)] dt = 0, \tag{1.8}$$

for all $\eta \in C^1([a,b];\mathbb{R}^n)$ such that $\eta(a) = \eta(b) = 0$.
Then

$$i) \quad g \in C^1([a,b];\mathbb{R}^n);$$

$$ii) \quad g'(t) = f(t), \qquad \forall t \in [a,b].$$

Proof– Suppose f is identically equal to 0. Recalling (1.8), for every constant c we have

$$\int_a^b (g(t) - c) \cdot \eta'(t)\ dt = \int_a^b g(t) \cdot \eta'(t)\ dt - c[\eta(b) - \eta(a)] = 0. \tag{1.9}$$

Let η be defined as

$$\eta(t) = \int_a^t (g(s) - \overline{c})\ ds$$

where the constant $\overline{c}$ is chosen to satisfy condition $\eta(b) = 0$. Obviously, the definition of η gives

$$\eta'(t) = g(t) - \overline{c}. \tag{1.10}$$

Since

$$\eta(b) = \int_a^b g(s)\ ds - (b-a)\overline{c},$$

we derive that

$$\overline{c} = \frac{1}{b-a} \int_a^b g(s)\ ds.$$

Therefore putting (1.10) into (1.9) we get

$$\int_a^b (g(t) - c) \cdot (g(t) - \overline{c})\ dt = 0, \qquad \forall c \in \mathbb{R}.$$

Choosing $c = \overline{c}$, the above equation can be rewritten as

$$\int_a^b (g(t) - \overline{c})^2\ dt = 0,$$

which yields $g(t) = \overline{c}$.
If f is not identically 0, we define a function F as

$$F(t) = \int_a^t f(s)\ ds.$$

Recalling (1.8) and integrating by parts we obtain

$$\begin{aligned} 0 &= \int_a^b [f(t)\cdot\eta(t) + g(t)\cdot\eta'(t)]\ dt = \int_a^b [F'(t)\cdot\eta(t) + g(t)\cdot\eta'(t)]\ dt \\ &= [F(b)\cdot\eta(b) - F(a)\cdot\eta(a)] + \int_a^b (g(t) - F(t))\cdot\eta'(t)\ dt \\ &= \int_a^b (g(t) - F(t))\cdot\eta'(t)\ dt. \end{aligned}$$

Now, exploiting the result obtained above, we have $g(t) - F(t) = k$, where k is a constant. Hence the lemma is proved. ■

We are now ready to state the *necessary conditions* that a minimizer ξ^* must satisfy.

Theorem 1.2.3 *Assume (1.4) and let ξ^* be a solution of (1.6). Then*

i) $\dfrac{\partial L}{\partial q}(\cdot, \xi^*(\cdot), \xi^{*\prime}(\cdot)) \in C^1([a,b]\times \mathbb{R}^n \times \mathbb{R}^n)$;

ii) $\dfrac{d}{dt}\dfrac{\partial L}{\partial q}(t, \xi^*(t), \xi^{*\prime}(t)) = \dfrac{\partial L}{\partial x}(t, \xi^*(t), \xi^{*\prime}(t)), \qquad \forall t \in [a,b].$ **(Euler equation)**

Proof– Consider any perturbation function $\eta : [a,b] \to \mathbb{R}^n$ such that $\eta \in C^1([a,b];\mathbb{R}^n)$ and $\eta(a) = \eta(b) = 0$. Then the arcs $\xi(t) = \xi^*(t) + \lambda\eta(t)$ are admissible for any constant λ. Since ξ^* is a solution of (1.6)

$$J(\xi^*) \le J(\xi^* + \lambda\eta), \qquad \lambda \in \mathbb{R}.$$

In other words, $J(\xi^* + \lambda\eta)$ has a minimum for $\lambda = 0$. Hence, applying the classic first order condition on J we get

$$\begin{aligned} 0 &= \frac{d}{d\lambda} J(\xi^* + \lambda\eta)|_{\lambda=0} \\ &= \int_a^b \left[< \frac{\partial L}{\partial x}(t, \xi^*(t), \xi^{*\prime}(t)), \eta(t) > + < \frac{\partial L}{\partial q}(t, \xi^*(t), \xi^{*\prime}(t)), \eta'(t) > \right] dt. \end{aligned}$$

The result follows from Lemma 1.2.2, taking

$$f(t) = \frac{\partial L}{\partial x}(t, \xi^*(t), \xi^{*\prime}(t)) \qquad \text{and} \qquad g(t) = \frac{\partial L}{\partial q}(t, \xi^*(t), \xi^{*\prime}(t)).$$

■

Remark 1.2.4 It is easy to prove that any local minimizer of problem (1.6) is a solution of the Euler equation. Indeed, if $\bar{\xi}$ is a local minimizer, then for any fixed $C^1([a,b], \mathbb{R}^n)$ function η such that $\eta(a) = \eta(b) = 0$ the map $g : \mathbb{R} \to \mathbb{R}$ defined by $g(h) := J(\bar{\xi} + h\eta)$ has a local minimum at $h = 0$. Hence $g'(0) = 0$ and the Euler equation derives from Lemma 1.2.2 as in the previous theorem.

Let us now consider another kind of problem, namely the Bolza problem

$$\min \left\{ \int_a^b L(t, \xi(t), \xi'(t))dt + \phi_a(\xi(a)) \mid \xi \in C^1([a,b]; \mathbb{R}^n), \xi(b) = \xi_b \right\}, \tag{1.11}$$

which is a problem with a free initial boundary condition, but not identically null initial cost.

Then, proceeding as we did for problem (1.6), we derive the Euler equation and an additional condition called *transversality condition*, which replaces the constraint on the initial state.

Theorem 1.2.5 *Assume (1.4) and let ξ^* be a solution of (1.11). Then*

i) $\displaystyle \frac{\partial L}{\partial q}(t, \xi^*, \xi^{*\prime}) \in C^1([a,b] \times \mathbb{R}^n \times \mathbb{R}^n);$

ii) $\displaystyle \frac{d}{dt}\frac{\partial L}{\partial q}(t, \xi^*(t), \xi^{*\prime}(t)) = \frac{\partial L}{\partial x}(t, \xi^*(t), \xi^{*\prime}(t)), \qquad \forall t \in [a,b];$

iii) $\displaystyle \frac{\partial L}{\partial q}(a, \xi^*(a), \xi^{*\prime}(a)) = D\phi_a(\xi^*(a)).$

Proof– Consider any perturbation function $\eta : [a,b] \to \mathbb{R}^n$ such that $\eta \in C^1([a,b]; \mathbb{R}^n)$ and $\eta(b) = 0$. Then the arcs $\xi(t) = \xi^*(t) + \lambda\eta(t)$ are admissible for any constant λ. Proceeding as in the previous theorem and integrating by parts, we have

$$\begin{aligned} 0 &= \frac{d}{d\lambda} J(\xi^* + \lambda\eta)|_{\lambda=0} \\ &= \int_a^b < \frac{\partial L}{\partial x}(t, \xi^*(t), \xi^{*\prime}(t)) - \frac{d}{dt}\frac{\partial L}{\partial q}(t, \xi^*(t), \xi^{*\prime}(t)), \eta(t) > dt \\ &+ < D\phi_a(\xi^*(a)) - \frac{\partial L}{\partial q}(a, \xi^*(a), \xi^{*\prime}(a)), \eta(a) >, \qquad \forall \eta. \end{aligned}$$

Hence the result follows. ∎

Remark 1.2.6 If we consider the case

$$\min \left\{ \int_a^b L(t, \xi(t), \xi'(t))\ dt + \phi_b(\xi(b)) : \xi \in C^1([a,b]; \mathbb{R}^n), \xi(a) = \xi_a \right\}, \tag{1.12}$$

then the transversality condition is

$$\frac{\partial L}{\partial q}(b, \xi^*(b), \xi^{*\prime}(b)) = -D\phi_b(\xi^*(b)),$$

as it can be easily seen by proceeding as in Theorem 1.2.5.

Definition 1.2.7 *The equation*

$$\frac{d}{dt}\frac{\partial L}{\partial q}(t, \xi^*(t), \xi^{*\prime}(t)) = \frac{\partial L}{\partial x}(t, \xi^*(t), \xi^{*\prime}(t)) \tag{1.13}$$

is called the Euler equation. *Any arc ξ^* which solves (1.13) is called an* extremal *for problem (1.6) or (1.11).*

The class of C^1 functions is not large enough to solve many problems in Calculus of Variations. This is the reason why we introduced the problem on the larger class of C^1_* functions. Although such a class is not the widest one on which we can set the problem it is general enough for the purpose of these notes. The next example shows a problem admitting C^1_* solutions, but no regular ones.

Example 1.2.8 Let $L : [-1, 1] \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ be defined as $L(t, x, q) = x^2(q-1)^2$, $S_{-1} = \{0\}$, $S_1 = \{1\}$ and $\phi_{-1} = \phi_1 = 0$. The problem is to minimize

$$\left\{\int_{-1}^{1} \xi(t)^2(\xi'(t)-1)^2 \, dt \mid \xi(-1) = 0, \xi(1) = 1\right\}.$$

It is easy to see that the problem, set either on the space $C^1([-1, 1]; \mathbb{R})$ or $C^1_*([-1, 1]; \mathbb{R})$, satisfies

$$\inf\left\{\int_{-1}^{1} \xi(t)^2(\xi'(t)-1)^2 \, dt \mid \xi \in A\right\} = 0,$$

since by definition

$$\inf\left\{\int_{-1}^{1} \xi(t)^2(\xi'(t)-1)^2 \, dt \mid \xi \in A\right\} \geq 0$$

and for the maps $\xi^*_n \in C^1([-1, 1]; \mathbb{R})$, $n \in \mathbb{N}$, given by

$$\xi^*_n(t) = \begin{cases} 0 & \text{for } t \in [-1, 0) \\ \frac{n}{2}t^2 & \text{for } t \in [0, 1/n) \\ t - \frac{1}{2n} & \text{for } t \in [1/n, 1] \end{cases}$$

we have

$$\begin{aligned} &\int_{-1}^{1} L(t, \xi^*(t), \xi^{*\prime}(t)) \, dt \\ &= \int_0^{1/n} \frac{n^2}{4} t^4(nt-1)^2 \, dt = \left[\frac{n^4t^7}{28} + \frac{n^2t^5}{20} - \frac{n^3t^6}{12}\right]_0^{1/n} \\ &= \frac{1}{28n^3} + \frac{1}{20n^3} - \frac{1}{12n^3} = \frac{1}{420n^3} \to 0, \qquad \text{as } n \to \infty. \end{aligned}$$

The minimum cannot be attained on the class of C^1 functions. Indeed, if a minimizer ξ^* exists, then by $\int_{-1}^{1} \xi^{*2}(t)(\xi^{*\prime}(t)-1)^2\, dt = 0$ we get that either $\xi^*(t) \equiv 0$ or $\xi^{*\prime}(t) \equiv 1$ on $[-1,1]$. But ξ^* has to verify also $\xi^*(-1)=0$, $\xi^*(1)=1$, which is impossible for any such a function. On the other hand, the minimum is achieved on the class of C^1_* functions, for example, by the map

$$\xi^*(t) = \begin{cases} 0 & \text{for } t \in [-1,0) \\ \\ t & \text{for } t \in [0,1]. \end{cases}$$

Let us go back to the minimization problems over the class of C^1_* functions, in particular to the fixed boundary conditions problem

$$\min \left\{ \int_a^b L(t,\xi(t),\xi'(t))\, dt \mid \xi \in C^1_*([a,b];\mathbb{R}^n), \xi(a)=\xi_a, \xi(b)=\xi_b \right\}, \tag{1.14}$$

and to the Bolza problem

$$\min \left\{ \int_a^b L(t,\xi(t),\xi'(t))\, dt + \phi_a(\xi(a)) \mid \xi \in C^1_*([a,b];\mathbb{R}^n), \xi(b)=\xi_b \right\}. \tag{1.15}$$

Let us see the appropriate formulation of the Euler equation for problems (1.14) and (1.15).

Theorem 1.2.9 *Assume (1.4) and let $\bar\xi$ be a local minimizer for problem (1.14) or (1.15). Then*

i) $\dfrac{\partial L}{\partial q}(\cdot,\bar\xi(\cdot),\bar\xi'(\cdot)) \in C^1_*([a,b]\times\mathbb{R}^n\times\mathbb{R}^n)$;

ii) there exists $c \in \mathbb{R}$ such that

$$\frac{\partial L}{\partial q}(t,\bar\xi(t),\bar\xi'(t)) = c + \int_a^t \frac{\partial L}{\partial x}(s,\bar\xi(s),\bar\xi'(s))\, ds, \qquad \forall t \in [a,b].$$

(Integral Euler equation)

Moreover, if $\bar\xi$ is a local minimizer for problem (1.15), then it also satisfies the transversality condition

$$\frac{\partial L}{\partial q}(a,\bar\xi(a),\bar\xi'(a)) = D\phi_a(\bar\xi(a)).$$

Proof– Let us start by considering problem (1.14). Take any function $\eta : [a,b] \to \mathbb{R}^n$ such that $\eta \in C^1_*([a,b];\mathbb{R}^n)$ and $\eta(a)=\eta(b)=0$. Then the arcs $\xi(t)=\bar\xi(t)+\lambda\eta(t)$ are admissible for any constant λ. Since $\bar\xi$ is a solution of (1.14)

$$J(\bar\xi) \le J(\bar\xi+\lambda\eta), \qquad \lambda \text{ small enough.}$$

In other words, $J(\bar{\xi} + \lambda\eta)$ has a local minimum for $\lambda = 0$. Hence, applying the classic first order condition on J we get

$$\begin{aligned} 0 &= \frac{d}{d\lambda} J(\bar{\xi} + \lambda\eta)|_{\lambda=0} \\ &= \sum_{i=1}^{N} \int_{t_{i-1}}^{t_i} \left[\frac{\partial L}{\partial x}(t, \bar{\xi}(t), \bar{\xi}'(t)) \cdot \eta(t) + \frac{\partial L}{\partial q}(t, \bar{\xi}(t), \bar{\xi}'(t)) \cdot \eta'(t) \right] dt, \end{aligned}$$

where $\Pi = \{t_0, t_1, \ldots, t_N\}$ is a common partition for $\bar{\xi}$ and η. As in Theorem 1.2.3 the result follows from Lemma 1.2.2, which holds true if we replace $C([a, b]; \mathbb{R}^n)$ by

$$C_*([a, b]; \mathbb{R}^n) = \left\{ \begin{array}{l} \xi \in C([t_i, t_{i+1}], \mathbb{R}^n),\ i = 0, \ldots, N-1, \text{ for some} \\ \text{partition } \Pi = \{a = t_0, t_1, \ldots, t_N = b\} \text{ of } [a, b] \end{array} \right\}$$

and $C^1([a, b]; \mathbb{R}^n)$ by $C^1_*([a, b]; \mathbb{R}^n)$. The integral Euler equation in the case of the Bolza problem (1.15) can be derived by the previous argument, provided we take into account the initial cost as in Theorem 1.2.5. Notice that any admissible function of the C^1_* case is continuously differentiable in a suitable right (left) neighbourhood of the initial (terminal) point a (b). ∎

Remark 1.2.10 Even though the proof of Theorem 1.2.3 shows that the Euler equation (1.13) holds true almost everywhere for problems set on the class of C^1_* functions, it is not the most suitable tool to provide a reasonable selection of candidate minimizers. Indeed, consider the problem

$$\min \left\{ \int_0^1 \xi'(t)^2 \, dt \mid \xi : [0, 1] \to \mathbb{R},\ \xi(0) = 0,\ \xi(1) = 1 \right\}. \tag{1.16}$$

Classical Euler equation (1.13) would give $\xi'' \equiv 0$ and then $\xi^*(t) = t$ in the case of minimization on C^1 functions. On the other hand, taking the minimization problem over the class of C^1_* functions and considering the equation (1.13) on subintervals, it yields an infinite number of extremals, i.e. all piecewise linear functions connecting the points $(0, 0)$ and $(1, 1)$ in the plane. Applying Theorem 1.2.9 instead of the classical Euler equation (1.13) we obtain that $\xi^*(t) = t$ is the unique candidate minimizer of the problem set over the class of piecewise C^1_* functions, as in the case of minimization on C^1 functions. As we will see in Section 1.2.2, this is due to the strict convexity of the Lagrangian with respect to the q variable.

Definition 1.2.11 *Any arc* $\bar{\xi} \in C^1_*([a, b]; \mathbb{R}^n)$ *which solves (for some* $c \in \mathbb{R}$*)*

$$\frac{\partial L}{\partial q}(t, \bar{\xi}(t), \bar{\xi}'(t)) = c + \int_a^t \frac{\partial L}{\partial x}(s, \bar{\xi}(s), \bar{\xi}'(s)) \, ds, \qquad \forall t \in [a, b].$$

is called an extremal *for problem (1.14) or (1.15).*

We conclude this section with a sufficient condition for the existence of global minimizers of problem (1.14).

Theorem 1.2.12 *Assume (1.4) and suppose that $L = L(t, x, q)$ is convex with respect to the pair of variables (x, q). Then, any extremal $\bar{\xi}$ of problem (1.14) which satisfies the boundary conditions $\bar{\xi}(a) = \xi_a$ and $\bar{\xi}(b) = \xi_b$ is a global minimizer.*

Proof– If $\bar{\xi}$ is an extremal, then there exists $c \in \mathbb{R}$ such that

$$\frac{\partial L}{\partial q}(t, \bar{\xi}(t), \bar{\xi}'(t)) = c + \int_a^t \frac{\partial L}{\partial x}(s, \bar{\xi}(s), \bar{\xi}'(s))\, ds, \qquad \forall t \in [a, b].$$

Setting $p(t) = \frac{\partial L}{\partial q}(t, \bar{\xi}^(t), \bar{\xi}'(t))$, we have

$$(p(t), p'(t)) = \left(\frac{\partial L}{\partial q}(t, \bar{\xi}(t), \bar{\xi}'(t)), \frac{\partial L}{\partial x}(t, \bar{\xi}(t), \bar{\xi}'(t))\right)$$

on any subinterval of continuity of $\bar{\xi}'$. Hence, due to the convexity[1] assumption on L, for any admissible ξ, we obtain

$$\begin{aligned}
& J(\xi) - J(\bar{\xi}) \\
= & \int_a^b \Big[L(t, \xi(t), \xi'(t)) - L(t, \bar{\xi}(t), \bar{\xi}'(t))\Big]\ dt \\
\geq & \int_a^b \left[\frac{\partial L}{\partial x}(t, \xi(t), \xi'(t)) \cdot (\xi(t) - \bar{\xi}(t)) + \frac{\partial L}{\partial q}(t, \bar{\xi}(t), \bar{\xi}'(t)) \cdot (\xi'(t) - \bar{\xi}'(t))\right]\ dt \\
= & \int_a^b \Big[p'(t) \cdot (\xi(t) - \bar{\xi}(t)) + p(t) \cdot (\xi'(t) - \bar{\xi}'(t))\Big]\ dt = \Big[p(t) \cdot (\xi(t) - \bar{\xi}(t))\Big]_a^b = 0.
\end{aligned}$$

■

1.2.2 Regularity and Erdmann condition

In this section we prove finer regularity properties of extremals for problem (1.14) under some injectivity assumptions on the Lagrangian, so that we can reformulate the minimization problem (1.14) as (1.6), restricting the set of admissible functions. Afterwards, we give an alternative formulation of the Euler equation in the case of autonomous problems, called *Erdmann condition.*

[1] We recall that whenever $f : \mathbb{R}^n \to \mathbb{R}$ is a convex function and x is a differentiability point of f, then

$$f(y) - f(x) \geq D_x f(x) \cdot (y - x) \qquad \forall y \in \mathbb{R}^n.$$

Proposition 1.2.13 *Let $\bar{\xi}$ be an extremal for problem (1.14) or (1.15) and suppose that for any $t \in [a,b]$ the map $q \in \mathbb{R}^n \mapsto \frac{\partial L}{\partial q}(t, \bar{\xi}(t), q)$ is injective. Then $\bar{\xi} \in C^1([a,b], \mathbb{R}^n)$ and satisfies the classical Euler Equation.*

Proof– Since $\bar{\xi} \in C^1_*([a,b]; \mathbb{R}^n)$, we already know that $\bar{\xi} \in C^1([t_i, t_{i+1}], \mathbb{R}^n)$ for any $i = 0, \ldots, N-1$. Let $\xi_i^+ = \lim_{t \to t_i^+} \bar{\xi}'(t)$ and $\xi_i^- = \lim_{t \to t_i^-} \bar{\xi}'(t)$. Since $\frac{\partial L}{\partial q}(\cdot, \bar{\xi}(\cdot), \bar{\xi}'(\cdot))$ is continuous by Theorem 1.2.9, then

$$\lim_{t \to t_i^{\pm}} \frac{\partial L}{\partial q}(t, \bar{\xi}(t), \bar{\xi}'(t)) = \frac{\partial L}{\partial q}(t, \bar{\xi}(t), \xi_i^{\pm}).$$

Hence, the injectivity property of $\frac{\partial L}{\partial q}(t, \bar{\xi}(t), q)$ implies that $\xi_i^+ = \xi_i^-$. ∎

Remark 1.2.14 Observe that $\frac{\partial L}{\partial q}(t, \bar{\xi}(t), q)$ is injective if, for example, $\frac{\partial^2 L}{\partial q^2}(t, \bar{\xi}(t), q) > 0$ for any $q \in \mathbb{R}^n$.

Remark 1.2.15 If $\frac{\partial L}{\partial q}(t, \bar{\xi}(t), q)$ is not injective, then problem (1.14) can admit solutions that are not C^1. Indeed, consider the problem

$$\min \left\{ \int_{-1}^{1} \xi'(t)^2 (t - \xi(t))^2 dt \mid \xi : [-1,1] \to \mathbb{R},\ \xi(-1) = 0,\ \xi(1) = 1 \right\}.$$

It is easy to see that such a problem is not solvable on $C^1([-1,1])$, while

$$\bar{\xi} = \begin{cases} 0 & \text{for } t \in [-1, 0), \\ t & \text{for } t \in [0, 1], \end{cases}$$

is a local minimizer in $C^1_*([-1,1])$.

Now, let us apply the previous regularity result to study a particular case of Bolza problem, whose Lagrangian depends on the q variable only. Indeed, the minimization problem (1.11) can be reduced to another minimization problem which is simpler to treat since it is set on a finite dimensional space. In what follows we will consider a free endpoint $t \in (a,b]$ and all possible terminal conditions $\xi(t) = x$, in order to study the behaviour of minimizers when t and x are varying.

The problem is to minimize for any $t \in (a,b]$ and $x \in \mathbb{R}^n$

$$J(\xi, t, x) = \left\{ \int_a^t L(\xi'(s))\ ds + \phi_a(\xi(a)) \mid \xi \in A_{t,x} \right\},$$

where $A_{t,x} = \{\xi \in C^1_*([a,b]; \mathbb{R}^n) \mid \xi(t) = x\}$. We denote by $v = v(t,x)$ the value of such a minimum (if it exists). Clearly $v(a,x) = \phi_a(x)$.

Theorem 1.2.16 *Assume that $L \in C^2(\mathbb{R}^n)$ is superlinear and strictly convex. Then Hopf's formula holds:*

$$v(t,x) = \min_{\xi \in A_{t,x}} J(\xi, t, x) = \min_{y \in \mathbb{R}^n} \left\{ (t-a) L\left(\frac{x-y}{t-a} \right) + \phi_a(y) \right\}. \tag{1.17}$$

Proof– First of all, Proposition 1.2.13 guarantees that we can restrict the problem over the class of C^1 functions. Moreover, since $L = L(q)$, the Euler equation and the transversality condition are respectively:

$$\frac{d}{dt}\frac{\partial L}{\partial q}(\xi^{*\prime}(t)) = 0,$$

$$\frac{\partial L}{\partial q}(\xi^{*\prime}(a)) = D\phi_a(\xi^*(a)).$$

Hence,

$$\frac{\partial L}{\partial q}(\xi^{*\prime}(s)) = D\phi_a(\xi^*(a)) \qquad \forall s \in [a,t].$$

Taking into account that $\frac{\partial^2 L}{\partial q^2} > 0$, we have

$$\xi^{*\prime}(s) = q^* = \frac{\partial L}{\partial q}^{-1}(D\phi_a(\xi^*(a)) \qquad \forall s \in [a,t].$$

Integrating the above equation and recalling that $\xi(t) = x$, we derive

$$\xi^*(s) = (s-t)q^* + x \qquad \forall s \in [a,t].$$

Setting $y^* = x - (t-a)q^*$ we get

$$v(t,x) = \int_a^t L(q^*)\ ds + \phi_a(x-(t-a)q^*) = (t-a)L\left(\frac{x-y^*}{t-a}\right) + \phi_a(y^*).$$

We have shown that

$$v(t,x) \geq \min_{y\in\mathbb{R}^n}\left\{(t-a)L\left(\frac{x-y}{t-a}\right) + \phi_a(y)\right\}.$$

On the other hand, to prove the opposite inequality, we recall that the Euler equation is

$$\xi^{*\prime\prime}(s) = 0 \qquad \forall s \in [a,t].$$

It follows that the minimizing arcs are straight lines and we can restrict the research of the minimum on linear admissible arcs $\xi(s) = (s-t)q + x$. Since v is the minimum of J:

$$v(t,x) \leq \int_a^t L(q)\ ds + \phi_a(x-(t-a)q) = (t-a)L\left(\frac{x-y}{t-a}\right) + \phi_a(y), \quad \forall y = x-(t-a)q.$$

The result follows passing to the minimum over $y \in \mathbb{R}^n$ in the above inequality. ∎

Example 1.2.17 Consider the case $n = 1$, $[a,b] = [0,1]$, $L(q) = q^2/2$ and

$$\phi_0(z) = \begin{cases} -z^2 & \text{if } |z| < 1 \\ 1 - 2|z| & \text{if } |z| \geq 1. \end{cases}$$

Using Hopf's formula we find that the minimum in (1.17) is attained at

$$\begin{cases} z = \dfrac{x}{1-2t} & \text{if } t < 1/2 \text{ and } |x| < 1-2t \\ \\ z = x + \text{sgn}(x)2t & \text{if } |x| \geq 1-2t \geq 0 \text{ or if } t \geq 1/2, \end{cases}$$

yielding

$$v(t,x) = \begin{cases} -\dfrac{x^2}{1-2t} & \text{if } t < 1/2 \text{ and } |x| < 1-2t \\ \\ 1 - 2(|x| + t) & \text{if } |x| \geq 1-2t \geq 0 \text{ or if } t \geq 1/2. \end{cases}$$

Definition 1.2.18 *The function $v : \mathbb{R}^+ \times \mathbb{R}^n \to \mathbb{R}$ defined by*

$$v(t,x) = \min_{\xi \in A_{t,x}} J(\xi, t, x)$$

is called value function.

Let's go back to the analysis of the regularity of extremals of problem (1.14) or (1.15).

Theorem 1.2.19 *Assume (1.4) and suppose that the Lagrangian L is of class C^k, $k \geq 2$, with respect to all variables and is strictly convex in the third variable, i.e. $\frac{\partial^2 L}{\partial q^2} > 0$. Then any minimizer ξ^* of J is of class $C^k([a,b], \mathbb{R}^n)$.*

Proof — We already know that $\xi^* \in C^1([a,b], \mathbb{R}^n)$; so it remains to show that if $\xi^* \in C^{h-1}([a,b], \mathbb{R}^n)$, $2 \leq h \leq k$, then $\xi^* \in C^h([a,b], \mathbb{R}^n)$. Let

$$P(t) = \int_a^t D_x L(r, \xi^*(r), \xi^{*\prime}(r)) \, dr$$

and

$$\Phi(t,q) = D_q L(t, \xi^*(t), q) - P(t), \qquad (t,q) \in [a,b] \times \mathbb{R}^n.$$

Then $\Phi \in C^{h-1}([a,b], \mathbb{R}^n)$ and the Euler equation gives

$$\Phi(t, \xi^{*\prime}(t)) = k_0 \qquad \forall t \in [a,b], \qquad \text{for some } k_0 \in \mathbb{R}.$$

Applying the Implicit Function Theorem to the map $F(t,q) = \Phi(t,q) - k_0$ in any point $(t_0, q_0) = (t_0, \xi^{*\prime}(t_0))$ (notice that $D_q F = \frac{\partial^2 L}{\partial q^2} > 0$), we easily obtain that $\xi^{*\prime}$ is a C^{h-1} function. ∎

Remark 1.2.20 Theorem 1.2.19 can be refined in the following sense. Assume (1.4) and suppose that along a minimizer ξ^* of J the following conditions hold:
i) the Lagrangian L is of class $C^k([a,b] \times U \times \mathbb{R}^n)$, for some $k \geq 2$, where U is an open bounded subset of $\mathbb{R}^n$ containing $\{\xi^*(t) : t \in [a,b]\}$;
ii) $\dfrac{\partial^2 L}{\partial q^2}(t, \xi^*(t), \xi^{*\prime}(t)) > 0$ for all $t \in [a,b]$.
Then ξ^* is of class $C^k([a,b], \mathbb{R}^n)$.

Now we are going to analyse a particular kind of problem, namely that having Lagrangian independent of t

$$L(t, x, q) = L(x, q).$$

For such a problem, called *autonomous problem*, the Euler equation yields the useful corollary known as the *Erdmann condition*:

Corollary 1.2.21 *Assume (1.4) and $L = L(x, q)$. Let ξ^* be an extremal of class C^2, i.e. a (regular) solution of the Euler equation. Then there exists a constant k such that*

$$L(\xi^*(t), \xi^{*\prime}(t)) - < \xi^{*\prime}(t), D_q L(\xi^*(t), \xi^{*\prime}(t)) >= k, \quad \textit{for all } t \in [a, b].$$

Proof– Exploiting the fact that ξ^* is a C^2 solution of the Euler equation, we obtain

$$\begin{aligned}
& \frac{d}{dt}[L(\xi^*(t), \xi^{*\prime}(t)) - < \xi^{*\prime}(t), D_q L(\xi^*(t), \xi^{*\prime}(t)) >] \\
= \;& < D_x L(\xi^*(t), \xi^{*\prime}(t)), \xi^{*\prime}(t) > + < D_q L(\xi^*(t), \xi^{*\prime}(t)), \xi^{*\prime\prime}(t) > \\
- \;& < \xi^{*\prime\prime}(t), D_q L(\xi^*(t), \xi^{*\prime}(t)) > - < \xi^{*\prime}(t), \frac{d}{dt} D_q L(\xi^*(t), \xi^{*\prime}(t)) > \\
= \;& < D_x L(\xi^*(t), \xi^{*\prime}(t)), \xi^{*\prime}(t) > - < \xi^{*\prime}(t), D_x L(\xi^*(t), \xi^{*\prime}(t)) >= 0
\end{aligned}$$

and the proof is complete. ■

1.2.3 Examples

In this section we will present three applications of Calculus of Variations to economics and we will also discuss some of the introductory examples (1.1.1-1.1.3-1.1.5). The economic minimization problem in Example 1.2.22 deals with a strictly convex Lagrangian with respect to the q variable, while the economic maximization problems 1.2.23 and 1.2.24 have a strictly concave Lagrangian. Hence, the regularity results of Section 1.2.2 apply, so that we can directly consider them on the space of C^1 functions. The Examples 1.1.1-1.1.3-1.1.5 were already defined on the class of C^1 functions and so they will be analysed in such a space. Let us start with the economic examples.

Example 1.2.22 Inflation and unemployment are the cause of social losses. The task is to find the optimal combination of inflation and unemployment under Philipps tradeoff in order to minimize the social loss function. The economic ideal is represented by the couple $(I_0, 0)$, where I_0 is the ideal income level when there is full employment and the ideal inflation rate is equal to 0.
Let $I(t)$ be the real income and $p(t)$ the inflation rate. Since any deviations from the ideal economic $(I_0, 0)$ is undesirable, we introduce the social loss function λ as follows:

$$\lambda(I, p) = (I_0 - I(t))^2 + \alpha p^2(t), \qquad \alpha > 0.$$

I and p are related through the expectation augmented Phillips tradeoff:

$$p(t) = -\beta(I_0 - I(t)) + \pi(t) \qquad \beta > 0, \tag{1.18}$$

where $I_0 - I(t)$ represents the shortfall of current national income from its full–employment level and π is the expected inflation rate which satisfies the adaptivity assumption:

$$\pi'(t) = \gamma[p(t) - \pi(t)] \qquad 0 < \gamma \leq 1, \quad \pi(0) = \pi_0, \quad \pi(T) = \pi_T \quad \pi_0, \pi_T \text{ assigned.} \tag{1.19}$$

The condition $\pi(T) = \pi_T$ represents a policy target.
The problem is the following: find $\pi(\cdot)$ that minimizes the discounted social loss over the time interval $[a, b]$, that is

$$\int_0^T e^{-\rho t} \lambda(I(t), p(t))\ dt \tag{1.20}$$

where $\rho > 0$ and $e^{-\rho t}$ is the discount factor. See [21] or [5, p.54] for further details.
From (1.18) and (1.19) it follows

$$I_0 - I(t) = \frac{\pi(t) - p(t)}{\beta} = -\frac{\pi'(t)}{\beta\gamma} \tag{1.21}$$

and again from (1.19) we have

$$p(t) = \frac{\pi'(t)}{\gamma} + \pi(t). \tag{1.22}$$

Using equalities (1.21) and (1.22), then functional (1.20) can be rewritten as functional $J[\pi]$ defined as

$$J[\pi] = \int_0^T e^{-\rho t} \Lambda(\pi(t), \pi'(t)) dt, \tag{1.23}$$

where

$$\Lambda(\pi, q) = \left[\frac{q}{\beta\gamma}\right]^2 + \alpha \left[\frac{q}{\gamma} + \pi\right]^2.$$

Hence the problem is to minimize J over all $\pi \in C^1([0, T]; \mathbb{R})$, such that $\pi(0) = \pi_0,\ \pi(T) = \pi_T$.
The Lagrange function L is the following

$$L(t, \pi, q) = e^{-\rho t} \left[\frac{q}{\beta\gamma}\right]^2 + \alpha \left[\frac{q}{\gamma} + \pi\right]^2,$$

and

$$\phi_0 = \phi_T = 0, \qquad S_0 = \{\pi_0\}, \qquad S_T = \{\pi_T\}.$$

Let us check if Theorem 1.1.11 can be applied to ensure the existence of minimizers.
Assumption (E1) holds since $\forall t$ such that $0 \leq t \leq T$

$$L(t, x, q) \geq e^{-\rho t} \left[\frac{q}{\beta\gamma}\right]^2 \geq e^{-\rho T} \left[\frac{q}{\beta\gamma}\right]^2,$$

and the Nagumo function θ can be defined as $\theta(r) = e^{-\rho T}\left[\frac{r}{\beta\gamma}\right]^2$.
As for assumption (E2) we remark that L is twice differentiable with respect to q and

$$\frac{\partial^2 L}{\partial q^2} = 2e^{-\rho t}\left[\frac{1}{(\beta\gamma)^2} + \frac{\alpha}{\gamma^2}\right] > 0.$$

Finally, $S_0 = \{\pi_0\}$ and $S_T = \{\pi_T\}$ are bounded. Hence all the hypotheses of Theorem 1.1.11 are satisfied and problem (1.20) has minimizers.
Let us now evaluate the candidate minimizers. The Euler equation

$$\frac{d}{dt}\frac{\partial L}{\partial q}(t, \pi(t), \pi'(t)) = \frac{\partial L}{\partial x}(t, \pi(t), \pi'(t))$$

associated with this problem is

$$\frac{d}{dt}\left\{2e^{-\rho t}\left[\frac{\pi'(t)}{(\beta\gamma)^2} + \frac{\alpha}{\gamma}(\frac{\pi'(t)}{\gamma} + \pi(t))\right]\right\} = 2\alpha e^{-\rho t}\left[\frac{\pi'(t)}{\gamma} + \pi(t)\right].$$

Hence, differentiating and performing several algebraic calculations, we get the boundary problem

$$\begin{cases} \pi''(t) - \rho\pi'(t) - \delta\pi(t) = 0 \\ \\ \pi(0) = \pi_0 \\ \\ \pi(T) = \pi_T, \end{cases}$$

where

$$\delta = \alpha\frac{(\beta\gamma)^2}{1 + \alpha\beta^2}\frac{\gamma + \rho}{\gamma}.$$

The general integral of the above second order linear homogeneous differential equation with constant coefficients is

$$\pi(t) = c_1 e^{r_+ t} + c_2 e^{r_- t}$$

where r_+, r_- are the characteristic roots

$$r_+ = \frac{\rho + \sqrt{\rho^2 + 4\delta}}{2} > 0$$

and

$$r_- = \frac{\rho - \sqrt{\rho^2 + 4\delta}}{2} < 0,$$

and c_1, c_2 are some constants to be evaluated exploiting the boundary conditions $\pi(0) = \pi_0$ and $\pi(T) = \pi_T$. In particular,

$$\begin{cases} c_1 + c_2 = \pi_0 \\ \\ e^{r_+ T} c_1 + e^{r_- T} c_2 = \pi_T, \end{cases}$$

from which we get

$$c_1 = \frac{\pi_0 e^{r_- T} - \pi_T}{e^{r_- T} - e^{r_+ T}} \qquad \text{and} \qquad c_2 = \frac{\pi_T - \pi_0 e^{r_+ T}}{e^{r_- T} - e^{r_+ T}}.$$

The optimal expected inflation rate π is then completely determined, since any minimizer of the minimization problem must be a solution of the Euler equation and we proved that such an equation admits a unique extremal.

Let us now consider the Hamermesh model of optimal adjustment of labor demand, which is an economic example with a fixed initial point, a variable terminal state and a terminal payoff.

Example 1.2.23 After a reduction of wages at $t = 0$, a firm wants to raise its labor input from the present state $\mathcal{L}_0$ to an undetermined one $\mathcal{L}(T) = \mathcal{L}_T$ in a fixed amount of time T. Adjusting labor input the firm faces a cost C which depends on the velocity of the adjustment: $C = C(\mathcal{L}'(t)) = a\mathcal{L}'(t)^2 + b$, where $a > 0$, and $b > 0$ if $\mathcal{L}'$ is not vanishing. The aim of the firm is to maximize the total discounted net profit Π over time. Let us denote by $\rho > 0$ the instantaneous interest rate. In the time interval that goes from 0 to T such a profit is equal to the definite integral between 0 and T of $[\pi(\mathcal{L}(t)) - C(\mathcal{L}'(t))]e^{-\rho t}$, where $\pi(\mathcal{L}(t))$ represents the profit of the firm at time t. In order to have a quantitative solution we choose $\pi(\mathcal{L}(t)) = 2c\mathcal{L}(t) - d\mathcal{L}^2(t)$, $0 < d < c$. For $t > T$ Π contains the term $\frac{1}{\rho}\pi(\mathcal{L}(T))e^{-\rho T}$ which represents the present value of a continuous perpetual rent of $\pi(\mathcal{L}_T)$ from time T onward. The problem is to maximize the functional

$$\Pi(\mathcal{L}) = \left\{ \int_0^T e^{-\rho t}[2c\mathcal{L}(t) - d\mathcal{L}^2(t) - a\mathcal{L}'(t)^2 - b]dt + \frac{1}{\rho}[2c\mathcal{L}(T) - d\mathcal{L}^2(T)]e^{-\rho T} \right\},$$

on functions $\mathcal{L} \in C^1([0,T];\mathbb{R}^n)$, $\mathcal{L}(0) = \mathcal{L}_0$ (see also [16] or [5, p. 75]).
In order to find the Euler equation of this problem we consider the Lagrange function associated with it and we evaluate its partial derivatives

$$L(t,x,q) = [2cx - dx^2 - aq^2 - b]e^{-\rho t},$$

$$L_x(t,x,q) = [2c - 2dx]e^{-\rho t},$$
$$L_q(t,x,q) = [-2aq]e^{-\rho t}.$$

Hence the Euler equation is

$$\frac{d}{dt}[-2a\mathcal{L}'(t)]e^{-\rho t} = 2[c - 2d\mathcal{L}(t)]e^{-\rho t}.$$

After several calculations this equation can be written as

$$\mathcal{L}''(t) - \rho\mathcal{L}'(t) - \frac{d}{a}\mathcal{L}(t) = -\frac{c}{a},$$

which is a second order non–homogeneous linear differential equation with constant coefficients. Its general integral is

$$\mathcal{L}(t) = C_1 e^{r_+ t} + C_2 e^{r_- t} + \mathcal{L}$$

where r_+, r_- are the characteristic roots

$$r_+ = \frac{\rho + \sqrt{\rho^2 + \frac{4d}{a}}}{2} > 0 \qquad \text{and} \qquad r_- = \frac{\rho - \sqrt{\rho^2 + \frac{4d}{a}}}{2} < 0,$$

$\mathcal{L}$ is the particular integral

$$\mathcal{L} = \frac{c}{d}$$

and C_1, C_2 are constants to be determined by exploiting the initial condition $\mathcal{L}(0) = \mathcal{L}_0$ and the transversality condition.
Hence

$$\mathcal{L}_0 = \mathcal{L}(0) = C_1 + C_2 + \mathcal{L},$$

which yields

$$C_1 + C_2 = \mathcal{L}_0 - \mathcal{L} =: D_1. \tag{1.24}$$

On the other hand, from the transversality condition

$$L_q(T, \mathcal{L}(T), \mathcal{L}'(T)) = -D\phi_T(\mathcal{L}(T)),$$

we get

$$-2a\mathcal{L}'(T)e^{-\rho T} = -\frac{1}{\rho}[2c - 2d\mathcal{L}(T)]e^{-\rho T}.$$

Substituting $\mathcal{L}(T) = C_1 e^{r_+ T} + C_2 e^{r_- T} + \mathcal{L}$ and $\mathcal{L}'(T) = C_1 r_+ e^{r_+ t} + C_2 r_- e^{r_- T}$ in the above equation, after several calculations we have

$$AC_1 + BC_2 = D_2, \tag{1.25}$$

where

$$A = (r_+ + \frac{d}{a\rho})e^{r_+ T}, \; B = (r_- + \frac{d}{a\rho})e^{r_- T} \text{ and } D_2 = \frac{c}{a\rho}.$$

Finally, recalling equations (1.24) and (1.25), we have

$$\begin{cases} C_1 + C_2 = D_1 \\ \\ AC_1 + BC_2 = D_2, \end{cases}$$

from which we get

$$C_1 = \frac{BD_1 - D_2}{B - A} \qquad \text{and} \qquad C_2 = \frac{D_2 - AD_1}{B - A}.$$

As in Example 1.2.22, it can be shown that the Lagrangian of this problem satisfies the assumptions of Tonelli's Theorem; hence the specific quantitative solution $\mathcal{L}(\cdot)$, determined by the above constants, is the unique maximizer.

We consider an economic example where we apply the Erdmann condition.

Example 1.2.24 In the Evans Model a monopolistic firm produces a single item with the following cost function

$$C(Q) = Q^2 + Q + 1, \tag{1.26}$$

where Q is the output produced by the firm, which we assume to be always equal to the quantity of the item the market requires. Q depends on the item price P and on the rate of variation P' of such a price:

$$Q = 1 - P + P'. \tag{1.27}$$

The firm's profit is

$$\pi = PQ - C, \tag{1.28}$$

hence substituting (1.26) and (1.27) into (1.28)

$$\pi(P, P') = -2P^2 + 3PP' + 4P - P'^2 - 3P' - 3.$$

The firm's aim is to maximize the total profit Π on the time interval $[0, T]$, i.e. to maximize

$$\Pi(P) = \int_0^T \pi(P(t), P'(t))\ dt$$

subject to

$$P(0) = P_0,\ P(T) = P_T,\ P_0 \neq P_T. \tag{1.29}$$

For further details on the model see also [11] or [5, p. 49]. Since $L(x, q) = -2x^2 + 4x - q^2 - 3q + 3xq - 3$ is a C^2, autonomous function, strictly concave in q, the Erdmann condition applies giving the following necessary condition for maximizers

$$P'(t)^2 + 4P(t) - 2P(t)^2 - 3 = K, \qquad \text{for some constant } K. \tag{1.30}$$

Differentiating (1.30) with respect to t, we get

$$2P'(t)[P''(t) - 2P(t) + 2] = 0. \tag{1.31}$$

Hence by (1.31) either $P'(t) = 0$, yielding $P(t) =$ constant (which is not admissible for conditions (1.29)), or

$$P''(t) - 2P(t) + 2 = 0.$$

In the latter case, the general solution is

$$P(t) = c_1 e^{\sqrt{2}t} + c_2 e^{-\sqrt{2}t} + 1, \tag{1.32}$$

where the constants c_1 and c_2 are determined exploiting (1.29), i.e. solving the linear system

$$\begin{cases} c_1 + c_2 + 1 = P(0) = P_0 \\ \\ e^{\sqrt{2}T} c_1 + e^{-\sqrt{2}T} c_2 = P(T) = P_T. \end{cases}$$

Hence

$$c_1 = \frac{1 - P_T + e^{-\sqrt{2}T}[1 - P_0]}{e^{-\sqrt{2}T} - e^{\sqrt{2}T}}$$

$$c_2 = \frac{P_T - 1 - e^{\sqrt{2}T}[1 - P_0]}{e^{-\sqrt{2}T} - e^{\sqrt{2}T}},$$

and substituting c_1 and c_2 into (1.32) we have the candidate maximizer P.

Let us now analyse Examples 1.1.1-1.1.3-1.1.5 by means of the previous theorems.

Example 1.2.25 Consider the functional

$$J(\xi) = \int_a^b \sqrt{1 + \xi'(t)^2}\, dt$$

arisen in Example 1.1.1. We want to minimize J over the class of $C^1([a,b])$ functions ξ such that $\xi(a) = A$ and $\xi(b) = B$. Since $L(q) = \sqrt{1+q^2}$ is a C^2, strictly convex and autonomous function, Theorem 1.2.19 and the Erdmann condition apply, giving the following necessary condition for minimizers:

$$\sqrt{1 + \xi'(t)^2} - \frac{\xi'(t)^2}{\sqrt{1 + \xi'(t)^2}} = k, \qquad \text{for some constant } k.$$

Hence

$$\sqrt{1 + \xi'(t)^2} = 1/k, \qquad \text{i.e.} \qquad \xi'(t) = \text{constant}.$$

Since minimizers must satisfy the boundary condition $\xi(a) = A$ and $\xi(b) = B$ we conclude that the unique solution to problem 1.1.1 is the line joining points (a, A) and (b, B).

Example 1.2.26 (The soap bubble problem) Consider the functional

$$J(\xi) = 2\pi \int_a^b \xi(x)\sqrt{1 + \xi'(x)^2}\, dx$$

from Example 1.1.3. We want to minimize it among all the regular functions ξ such that $\xi(a) = A$ and $\xi(b) = B$ $(A, B \neq 0)$. In this case $L(y, q) = y\sqrt{1+q^2}$ is C^2, autonomous, but not strictly convex with respect to q, since $\frac{\partial^2 L}{\partial q^2}(y, q) = y(1+q^2)^{-3/2}$ is zero whenever y is. Anyway, since $A, B \neq 0$, any minimizer of J cannot be identically zero on $[a, b]$ and it can be shown that for suitable choices of a, b, A and B the Erdmann condition applies, giving

$$\xi(x)\sqrt{1 + \xi'(x)^2} - \frac{\xi(x)\xi'(x)^2}{\sqrt{1 + \xi'(x)^2}} = k \qquad \text{for some} \qquad k \in \mathbb{R}.$$

Hence

$$\frac{\xi(x)}{\sqrt{1 + \xi'(x)^2}} = k, \qquad \text{i.e.} \qquad \xi'(x)^2 = \frac{\xi(x)^2}{k^2} - 1.$$

This equation is solvable in any nonincreasing (or nondecreasing) interval of the extremals, with solutions of the following type

$$\xi(x) = \pm k \cosh\left(\frac{x - x_0}{k}\right)$$

(x_0 depending on the integration interval). Therefore minimizers of J must be locally arcs of hyperbolic functions.

Example 1.2.27 Consider the functional

$$J(\xi) = \frac{1}{\sqrt{2g}} \int_a^b \frac{1}{\sqrt{H - \xi(x)}} \sqrt{1 + \xi'(x)^2}\, dx$$

arisen in the Brachistochrone Problem (see Example 1.1.5). We want to minimize J over the class of $C^1([a,b])$ functions ξ such that $\xi(a) = A$ and $\xi(b) = B$, where $a < b$ and $A > B$. By physical reasonings we can suppose that minimizers ξ of J are nonincreasing functions, so that we can study $L(y,q) = \sqrt{\frac{1+q^2}{H-y}}$ (remember that $H = A + \frac{v_0^2}{2g}$, where $v_0 > 0$) in the closed strip $[m, M] \times \mathbb{R}$, where m, M are suitable values such that $m < B < A < M$. In this strip L is C^2 with respect to all variables and strictly convex in q; hence any minimizer ξ of J is a C^2 function that satisfies the Erdmann condition. Therefore ξ solves

$$\sqrt{\frac{1 + \xi'(x)^2}{H - \xi(x)}} - \frac{\xi'(x)^2}{\sqrt{(1 + \xi(x)^2)(H - \xi(x))}} = k, \qquad \text{for some } k \in \mathbb{R},$$

i.e. $(1 + \xi(x)^2)(H - \xi(x)) = 2c$, where $2c = 1/k^2 > 0$. Now, if we change variables, parameterizing the curve $\{(x, \xi(x)) : x \in [a,b]\}$ as $\{(x(t), \xi(t)) : t \in [\tau_1, \tau_2]\}$, where the parameter t satisfies $\frac{d\xi}{dx}(x(t)) = -\cot(t/2)$, we obtain that $x(t) = \tilde{c} + c(t - \sin t)$, $\xi(t) = H - c(1 - \cos t)$ for some $\tilde{c} \in \mathbb{R}$ and $t \in [\tau_1, \tau_2]$. Finally, c, $\tilde{c}$, τ_1 and τ_2 are uniquely determined imposing conditions $x(\tau_1) = a$, $x(\tau_2) = b$, $\xi(\tau_1) = A$ and $\xi(\tau_2) = B$.

1.2.4 Hamilton equations

In the previous sections we have proved that any minimizing arc of the functional J is also an extremal. This property is usually not enough in practice to determine minimizers. A first reason is that being an extremal is only a necessary condition for being a minimizer. Moreover, finding all extremals may not be an easy task, so it is important to have additional tools to study minimizers. This will be done in the case of strictly convex Lagrangian, which ensures at least C^1 regularity of extremals.

So let us consider again the Bolza problem of minimization of the functional

$$J(\xi) = \int_a^b L(t, \xi(t), \xi'(t))\, dt + \phi_a(\xi(a)) \tag{1.33}$$

over the class

$$A = \{\xi \in C^1([a,b]; \mathbb{R}^n) : \xi(b) = \xi_b\},$$

where $a, b \in \mathbb{R}$, $a < b$, are fixed and $L : [a, b] \times \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$, $\phi_a : \mathbb{R}^n \to \mathbb{R}$ are given continuous functions.
Let us set

$$H(t, x, p) = \sup_{q \in \mathbb{R}^n} [p \cdot q - L(t, x, q)]. \tag{1.34}$$

The function H is the Legendre–Fenchel transform of L with respect to the third argument (see Appendix A) and is called the *Hamiltonian* associated with L.

Euler's equation (1.13) can be equivalently restated as a first order system of $2n$ equations, as it is shown below. Such a formulation, called *Hamiltonian*, is more convenient to use in many cases.

Theorem 1.2.28 *Assume that $L \in C^2([a, b] \times \mathbb{R}^n \times \mathbb{R}^n)$ is superlinear and strictly convex with respect to the third variable. Let $\xi^* \in C^2([a, b], \mathbb{R}^n)$ be an extremal for problem (1.33) and let us set*

$$\eta^*(t) = D_q L(t, \xi^*(t), \xi^{*\prime}(t)), \qquad t \in [a, b].$$

Then $\eta^*(a) = D\phi_a(\xi^*(a))$ and the pair (ξ^*, η^*) satisfies

$$\begin{cases} \xi^{*\prime}(t) = D_p H(t, \xi^*(t), \eta^*(t)) \\ \\ \eta^{*\prime}(t) = -D_x H(t, \xi^*(t), \eta^*(t)). \end{cases} \tag{1.35}$$

Conversely, suppose that $\xi^*, \eta^* \in C^2([a, b], \mathbb{R}^n)$ solve system (1.35) *together with the condition $\eta^*(a) = D\phi_a(\xi^*(a))$. Then ξ^* is an extremal for problem (1.33).*

Proof — Since $D_q L(t, x, \cdot)$ and $D_p H(t, x, \cdot)$ are reciprocal inverse (see (3.10)), the definition of η^* implies that

$$\xi^{*\prime}(t) = D_p H(t, \xi^*(t), \eta^*(t)) \qquad t \in [a, b].$$

On the other hand, since ξ^* is an extremal, we have

$$\eta^{*\prime}(t) = D_x L(t, \xi^*(t), \xi^{*\prime}(t)) \qquad \eta^*(a) = D\phi_a(\xi^*(a)). \tag{1.36}$$

Recalling that $-D_x H(t, x, D_q L(t, x, q)) = D_x L(t, x, q)$ (see (3.7)) we obtain the first part of the assertion. The converse part is obtained by similar arguments. ∎

Example 1.2.29 Consider the map $L : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$, defined by

$$L(x, q) = ax^2 + bq^2, \qquad \text{with} \qquad b > 0,$$

and the minimization of functional

$$J(\xi) = \int_0^1 L(\xi(t), \xi'(t))\, dt$$

over the class $A = \{\xi \in C^1([0,1];\mathbb{R}^n) : \xi(0) = x_0, \xi(1) = x_1\}$. Then

$$H(t,p) = \sup_{q\in\mathbb{R}}(p\cdot q - ax^2 - bq^2) = -ax^2 + \sup_{q\in\mathbb{R}^n}(p\cdot q - bq^2) = -ax^2 + \frac{p^2}{4b}$$

and system (1.35) becomes

$$\begin{cases} \xi'(t) = \dfrac{\eta(t)}{2b}, \\ \\ \eta'(t) = 2a\xi(t). \end{cases}$$

If we differentiate the first equation we obtain $\xi''(t) = \frac{\eta'(t)}{2b} = \frac{a\xi(t)}{b}$. Let us distinguish the following cases.

1. Case $a > 0$.
   If we set $\lambda = \sqrt{\frac{a}{b}}$, then solutions to $\xi''(t) = \frac{a\xi(t)}{b}$ are given by

   $$\xi(t) = c_0 e^{\lambda t} + c_1 e^{-\lambda t},$$

   where $c_0, c_1 \in \mathbb{R}$ are uniquely determined by the boundary conditions. Hence in this case our problem admits a unique extremal, given by

   $$\xi(t) = \frac{1}{e^{-\lambda} - e^{\lambda}}\Big((x_0 e^{-\lambda} - x_1)e^{\lambda t} + (x_1 - x_0 e^{\lambda})e^{-\lambda t}\Big).$$

2. Case $a = 0$.
   All the solutions of $\xi''(t) = 0$ are straight lines. Then the unique extremal of the problem is $\xi(t) = x_0 + (x_1 - x_0)t$.

3. Case $a < 0$.
   If we set $\lambda = \sqrt{\frac{|a|}{b}}$, then solutions to $\xi''(t) = \frac{a\xi(t)}{b}$ are given by

   $$\xi(t) = c_0 \sin\lambda t + c_1 \cos\lambda t,$$

   where $c_0, c_1 \in \mathbb{R}$ depend on the boundary conditions. But in this case we do not always have uniqueness of solutions. Indeed, if we consider $x_0 = x_1 = 0$, $a = -\pi^2$, $b = 1$, then the map $\xi(t) = c_0 \sin\lambda t$ is an extremal for the functional $J(\xi)$ for any choice of $c_0 \in \mathbb{R}$.

1.2.5 Exercises

Exercise 1.2.30 *Find all the extremals for the functionals J given below.*

- $J(\xi) = \displaystyle\int_0^1 [t + 2\xi(t) + \frac{1}{2}\xi'(t)]^2\, dt;$
- $J(\xi) = \displaystyle\int_0^1 e^t[\xi(t)^2 + \frac{1}{2}\xi'(t)]^2\, dt;$

- $J(\xi) = \int_0^1 [\xi'_1(t)^2 + \xi'_2(t)^2 - 2\xi_1(t)\xi_2(t)]\ dt, \qquad \xi = (\xi_1, \xi_2).$

Exercise 1.2.31 *Fix $T > 0$ and $x_0 \in \mathbb{R}^n$. Minimize functional*

$$J(\xi) = \int_0^T [2|\xi(s)|^2 + \frac{1}{2}|\xi'(s)|^2]\ ds + \frac{1}{2}|\xi(0)|^2$$

over the class of C^1 functions $\xi : [0,T] \to \mathbb{R}^n$ such that $\xi(T) = x_0$.

Exercise 1.2.32 *Minimize functional*

$$J(\xi) = \int_0^1 \left[\frac{1}{2}\xi'(s)^2 + \frac{1}{2}\xi(s)^2 - \xi(s)\right]\ ds + \xi(0)^2$$

over the class of C^1 functions $\xi : [0,1] \to \mathbb{R}$ such that $\xi(1) = 1$.

Exercise 1.2.33 *Minimize functional*

$$J(\xi) = \int_0^1 \left[\frac{1}{2}\xi_1'(s)^2 + \frac{1}{2}\xi_2'(s)^2 + \frac{1}{4}(\xi_1(s) + \xi_2(s))^2\right] ds + \xi_1(0) - \xi_2(0)$$

over the class of C^1 functions $\xi = (\xi_1, \xi_2) : [0,1] \to \mathbb{R}^2$ such that $(\xi_1(1), \xi_2(1)) = (1,1)$.

Exercise 1.2.34 *Consider the following functional and boundary conditions:*

$$J(\xi) = \int_0^1 (1+t)(\xi'(t)^2 + \xi(t))\ dt \qquad \xi(0) = 0,\ \xi(1) = 1,$$

where J is defined on the class of C^1 functions $\xi : [0,1] \to \mathbb{R}$. Find the extremals of J satisfying the above boundary conditions.

1.3 Second Order Conditions

Let us consider again the problem with fixed extremes:

$$\text{to minimize } \left\{\int_a^b L(t, \xi(t), \xi'(t))dt \mid \xi \in C^1_*([a,b];\mathbb{R}^n), \xi(a) = \xi_a, \xi(b) = \xi_b\right\}. \qquad (1.14)$$

Necessary conditions for minimizers of problem (1.14) arisen in section 1.2 are called *first order conditions* because they all come from the analysis of the *first variation* of L, that is of

$$\frac{d}{d\lambda} J(\xi^* + \lambda\eta)|_{\lambda=0},$$

where $J(\xi) = \int_a^b L(t, \xi(t), \xi'(t))dt$, ξ^* is a minimizer of J and η is any fixed C^1_* function such that $\eta(a) = 0$ and $\eta(b) = 0$. Existence of such a derivative is guaranteed by the assumption $L = L(t, x, q)$ is of class $C^1([a,b] \times \mathbb{R}^n \times \mathbb{R}^n)$.

If we assume $L \in C^2([a,b] \times \mathbb{R}^n \times \mathbb{R}^n)$ a second order necessary condition can be stated, called the *Legendre condition*.

Let us analyse the map $g : \mathbb{R} \to \mathbb{R}$ defined by $g(h) := J(\bar{\xi} + h\eta)$ when L is of class C^2. Notice that g has the same regularity as the Lagrangian L. So, in this case g is a C^2 map. Since

$$g'(h) = \int_a^b \Big[L_x(t, \bar{\xi}(t) + h\eta(t), \bar{\xi}'(t) + h\eta'(t)) \cdot \eta(t) + L_q(t, \bar{\xi}(t) + h\eta(t), \bar{\xi}'(t) + h\eta'(t)) \cdot \eta'(t) \Big] \, dt,$$

we obtain

$$g''(0) = \int_a^b [L_{xx}(t, \bar{\xi}(t), \bar{\xi}'(t))\eta(t) \cdot \eta(t) + 2L_{xq}(t, \bar{\xi}(t), \bar{\xi}'(t))\eta'(t) \cdot \eta(t) + L_{qq}(t, \bar{\xi}(t), \bar{\xi}'(t))\eta'(t) \cdot \eta'(t)] \, dt.$$

The fact that $\bar{\xi}$ is a local minimizer implies that $g''(0) \geq 0$. Now, we are ready to state the *Legendre condition*, which was first proved by Legendre in the second half of the seventeenth century.

Lemma 1.3.1 (Legendre condition) *Assume that $L = L(t, x, q)$ is of class $C^2([a,b] \times \mathbb{R}^n \times \mathbb{R}^n)$ and let $\bar{\xi}$ be a local minimizer of problem (1.14). Then, for any $t \in [a,b]$ the matrix*

$$P(t) := L_{qq}(t, \bar{\xi}(t), \bar{\xi}'(t)) \quad \textit{is positive semidefinite.} \tag{1.37}$$

Proof– Suppose first that $n = 1$. Then $P(t)$ is one-dimensional and assumption (1.37) reads $P(t) \geq 0$ for any $t \in [a,b]$. So, suppose by contradiction that there exists $t_0 \in [a,b]$ such that $P(t_0) < 0$. If $t_0 \in (a,b)$ and t_0 is a continuity point of $\bar{\xi}'$, we can find an open neighbourhood of t_0, say $\{t : |t - t_0| < \delta\}$ where $P(\cdot)$ is continuous and $P(t) < -\varepsilon$, for some positive ε. Moreover, for any $k \in \mathbb{N}$ we can find $n_k \in \mathbb{N}$ such that $\frac{n_k}{k} \leq t_0 < \frac{n_k+1}{k}$. For any k sufficiently large so that $\left[\frac{n_k}{k}, \frac{n_k+1}{k}\right] \subset (t_0 - \delta, t_0 + \delta)$, consider the C^1 maps

$$\eta_k(t) = \begin{cases} 0 & \text{for } t \in [a,b] \setminus \left[\frac{n_k}{k}, \frac{n_k+1}{k}\right), \\ \sin^2(k\pi t) & \text{for } t \in \left[\frac{n_k}{k}, \frac{n_k+1}{k}\right). \end{cases}$$

We have

$$\left| \int_a^b [L_{xx}(t, \bar{\xi}(t), \bar{\xi}'(t))\eta_k^2(t) + 2L_{xq}(t, \bar{\xi}(t), \bar{\xi}'(t))\eta_k'(t)\eta_k(t)] \, dt \right|$$

$$= \left| \int_{\frac{n_k}{k}}^{\frac{n_k+1}{k}} [L_{xx}(t, \bar{\xi}(t), \bar{\xi}'(t))\eta_k^2(t) + 2L_{xq}(t, \bar{\xi}(t), \bar{\xi}'(t))\eta_k'(t)\eta_k(t)] \, dt \right|$$

$$\leq \frac{1}{k}\|\bar{L}_{xx}\|_\infty^* + 2(2k\pi)\frac{1}{k}\|\bar{L}_{xq}\|_\infty^*,$$

where

$$\|\bar{L}_{xx}\|_\infty^* := \max_{i=1,\dots,N} \max_{t\in[t_i,t_{i+1}]} |L_{xx}(t,\bar{\xi}(t),\bar{\xi}'(t))|$$

and

$$\|\bar{L}_{xq}\|_\infty^* := \max_{i=1,\dots,N} \max_{t\in[t_i,t_{i+1}]} |L_{xq}(t,\bar{\xi}(t),\bar{\xi}'(t))|$$

(as usual, $\Pi = \{t_0, t_1, \dots, t_N\}$ is the partition $[a,b]$ that appears in Definition 1.1.6). On the other hand,

$$\begin{aligned}\int_a^b P(t)\eta_k'^2(t)\ dt &= \int_{\frac{n_k}{k}}^{\frac{n_k+1}{k}} P(t)\eta_k'^2(t)\ dt < -\varepsilon \int_{\frac{n_k}{k}}^{\frac{n_k+1}{k}} k^2\pi^2 \sin^2(2k\pi t)\ dt \\ &= -\varepsilon \int_{\frac{n_k}{k}}^{\frac{n_k+1}{k}} k^2\pi^2 \frac{1-\cos(4k\pi t)}{2}\ dt = -\varepsilon k^2\pi^2 \frac{1}{2k}.\end{aligned}$$

Then we have

$$\int_a^b [L_{xx}(t,\bar{\xi}(t),\bar{\xi}'(t))\eta_k^2(t) + 2L_{xq}(t,\bar{\xi}(t),\bar{\xi}'(t))\eta_k'(t)\eta_k(t) + L_{qq}(t,\bar{\xi}(t),\bar{\xi}'(t))\eta_k'^2(t)]\ dt$$

$$\leq \frac{1}{k}\|\bar{L}_{xx}\|_\infty^* + 4\pi\|\bar{L}_{xq}\|_\infty^* - \frac{\varepsilon k\pi^2}{2} < 0, \text{ for any } k \text{ sufficiently large,}$$

against the fact that $g''(0) \geq 0$ for any choice of $\eta \in C_*^1([a,b])$ such that $\eta(a) = \eta(b) = 0$. If $t_0 = a, b$ or t_0 is a discontinuity point of $\bar{\xi}'$, the same argument applies replacing t_0 with some t_0' close to t_0 where we still have $P(t_0') < 0$.
In the n-dimensional case we can proceed as follows. Suppose again by contradiction that $P(t_0) < 0$ for some $t_0 \in [a,b]$, i.e. the matrix $P(t_0)$ is negative definite. Let λ be a negative eigenvalue of $P(t_0)$ and θ the corresponding eigenvector. Consider the functions $\bar{\eta}_k(t) = \eta_k(t)\theta$, $k \in \mathbb{N}$, where η_k are the scalar functions defined above. Repeating the above argument with $\bar{\eta}_k$ and taking into account the local continuity of the map $t \mapsto P(t)\theta\cdot\theta$, we can derive a contradiction as before. ∎

Notice that the previous theorem holds true also for problem (1.6), as the proof shows.

Example 1.3.2 (Wirtinger) The Legendre condition is only a necessary condition. Indeed, consider the problem

$$\min \left\{ \int_0^T (\xi'(t)^2 - \xi(t)^2)\ dt \mid \xi \in C^1([0,T]), \xi(0) = \xi(T) = 0 \right\}.$$

Here the Lagrangian $L(x,q) = q^2 - x^2$ satisfies the Legendre condition and the arc $\bar{\xi} \equiv 0$ is an extremal. But, depending on the value of T, $\bar{\xi}$ can either be or not a local minimizer of functional $J(\xi) = \int_0^T (\xi'(t)^2 - \xi(t)^2)dt$. In particular, if $T \leq 2\sqrt{2}$, then $\bar{\xi}$ is a global minimizer, since $J(\bar{\xi}) = 0$ and for any other $\xi \in C^1([0,T])$ such that $\xi(0) = \xi(T) = 0$ we have (by the Holder inequality[2])

$$|\xi(t)| \le \int_0^t |\xi'(s)|\, ds \le \left(\int_0^t \xi'(s)^2\, ds\right)^{1/2} t^{1/2}$$

and

$$|\xi(t)| \le \int_t^T |\xi'(s)|\, ds \le \left(\int_t^T \xi'(s)^2\, ds\right)^{1/2} (T-t)^{1/2}.$$

Hence,

$$\int_0^{T/2} \xi(t)^2\, dt \le \int_0^{T/2} \left(\int_0^{T/2} \xi'(s)^2\, ds\right) t\, dt = \frac{T^2}{8} \int_0^{T/2} \xi'(s)^2\, ds$$

and

$$\int_{T/2}^T \xi(t)^2\, dt \le \int_{T/2}^T \left(\int_{T/2}^T \xi'(s)^2\, ds\right) (T-t)\, dt = \frac{T^2}{8} \int_{T/2}^T \xi'(s)^2\, ds.$$

Adding the above inequalities we deduce that

$$\int_0^T \xi(t)^2\, dt \le \int_0^T \xi'(t)^2\, dt,$$

that is $J(\xi) \ge 0$. But for $T > \sqrt{10}$ it is possible to find arcs ξ such that $\max_{t\in[0,T]} |\xi(t)|$, $\max_{t\in[0,T]} |\xi'(t)|$ are sufficiently small and $J(\xi) < 0$. For example, for any $n \in \mathbb{N}$ big enough the functions $\xi_n(t) = \frac{1}{n} t(T-t)$ are close to the 0 function in the C^1 norm and verify $J(\xi_n) = \frac{T^3}{3n^2}\left(1 - \frac{T^2}{10}\right) < 0$.

The previous example shows that even if L and the extremal of J are C^∞ functions and L is strictly convex, the Legendre condition is not sufficient to determine minimizers of J. However, Legendre thought that the following assertion was true:

False Theorem of Legendre

If $\bar{\xi}$ is a C^2 extremal of problem (1.14), where L is a $C^3([a,b] \times \mathbb{R}^n \times \mathbb{R}^n)$ Lagrangian and if $L_{qq}(t, \bar{\xi}(t), \bar{\xi}'(t)) > 0$ for any $t \in [a,b]$, then $\bar{\xi}$ is a local minimizer for functional J in (1.14).

It is important to see the proof given by Legendre, although false, because Jacobi could later state a *sufficient condition* for extremal to be minimizers by rectifying it.

[2] Holder Inequality for continuous scalar functions is the following: for $f, g \in C([a,b])$,

$$\int_a^b f(t)g(t)dt \le \left(\int_a^b f^2(t)dt\right)^{1/2} \left(\int_a^b g^2(t)dt\right)^{1/2}.$$

It can be proved as a consequence of the following inequality: for any $\alpha > 0$

$$f(t)g(t) \le \frac{\alpha}{2} f^2(t) + \frac{1}{2\alpha} g^2(t).$$

Proof– We will give the "proof" only in the one-dimensional case. For any fixed $C^1([a,b])$ function η such that $\eta(a) = \eta(b) = 0$ consider the map $g : \mathbb{R} \to \mathbb{R}$ defined by $g(h) := J(\bar\xi + h\eta)$. Since $\bar\xi$ is an extremal, then $g'(0) = 0$. Moreover, since L is C^3 and $\bar\xi$ is C^2,

$$\begin{aligned} g''(0) &= \int_a^b \Big[L_{xx}(t,\bar\xi(t),\bar\xi'(t))\eta^2(t) + L_{qq}(t,\bar\xi(t),\bar\xi'(t))\eta'^2(t) \\ &\quad + 2L_{xq}(t,\bar\xi(t),\bar\xi'(t))\eta'(t)\eta(t)\Big]\, dt \\ &= \int_a^b \Big[L_{xx}(t,\bar\xi(t),\bar\xi'(t))\eta(t)^2 + L_{qq}(t,\bar\xi(t),\bar\xi'(t))\eta'^2(t) \\ &\quad + \frac{d}{dt}\{L_{xq}(t,\bar\xi(t),\bar\xi'(t))\eta^2(t)\} - \frac{d}{dt}\{L_{xq}(t,\bar\xi(t),\bar\xi'(t))\}\eta^2(t)\Big]\, dt \\ &= \int_a^b \Big[(L_{xx}(t,\bar\xi(t),\bar\xi'(t)) - \frac{d}{dt}\{L_{xq}(t,\bar\xi(t),\bar\xi'(t))\})\eta(t)^2 \\ &\quad + L_{qq}(t,\bar\xi(t),\bar\xi'(t))\eta'^2(t)\Big]\, dt = \int_a^b \Big[Q(t)\eta(t)^2 + P(t)\eta'^2(t)\Big] dt, \end{aligned}$$

where

$$Q(t) := L_{xx}(t,\bar\xi(t),\bar\xi'(t)) - \frac{d}{dt}\{L_{xq}(t,\bar\xi(t),\bar\xi'(t))\}$$

and

$$P(t) := L_{qq}(t,\bar\xi(t),\bar\xi'(t)).$$

We would like to prove that $g''(0) > 0$ for any $\eta \neq 0$. Since $P(t) > 0$, for any C^1 function $R(t)$ we have

$$\begin{aligned} g''(0) &= \int_a^b \Big[Q(t)\eta(t)^2 + P(t)\eta'^2(t)\Big]\, dt \\ &= \int_a^b \Big[Q(t)\eta(t)^2 + P(t)\eta'^2(t) + \frac{d}{dt}(R(t)\eta^2(t))\Big]\, dt \\ &= \int_a^b \Big[Q(t)\eta(t)^2 + P(t)\eta'^2(t) + R'(t)\eta^2(t) + 2R(t)\eta(t)\eta'(t)\Big]\, dt \\ &= \int_a^b P(t)\Big[\eta'^2(t) + \frac{Q(t)+R'(t)}{P(t)}\eta(t)^2 + 2\frac{R(t)}{P(t)}\eta(t)\eta'(t)\Big]\, dt. \end{aligned}$$

Hence if we can prove that there exists a $C^1([a,b])$ function $R(t)$ such that

$$\frac{Q(t)+R'(t)}{P(t)} = \frac{R^2(t)}{P^2(t)}, \qquad \text{i.e.} \qquad R'(t) = \frac{R^2(t)}{P(t)} - Q(t),$$

we obtain

$$g''(0) = \int_a^b P(t)\Big\{\Big(\eta'^2(t) + \frac{R(t)}{P(t)}\eta(t)\Big)^2\Big\}\, dt.$$

Of course, if such a function exists, then $g''(0) > 0$ for any $\eta \neq 0$ and the proof of the theorem is "almost" finished[3]. Unfortunately, the previous equation does not always admit global solutions, as Legendre erroneously stated. However, on the basis of the above argument, Jacobi could prove another sufficient condition for the existence of local minimizers .

Jacobi's idea was to better analyse equation

$$R'(t) = \frac{R^2(t)}{P(t)} - Q(t), \tag{1.38}$$

trying to find conditions that could guarantee global existence of solutions. He looked for solutions of the following kind:

$$R(t) = -P(t)\frac{\eta'(t)}{\eta(t)}, \qquad \text{where} \qquad \eta(t) \neq 0,\ \forall t \in [a,b].$$

Substituting $R(t)$ as above in (1.38) we obtain the equation

$$-\frac{d}{dt}(P(t)\eta'(t)) + Q(t)\eta(t) = 0, \quad t \in [a,b] \quad \textbf{(Jacobi equation)}. \tag{1.39}$$

Before stating Jacobi's result, let us introduce the following notion:

Definition 1.3.3 *A point $c \in (a,b]$ is called a* conjugate point *(to a) if there exists a nontrivial solution η of (1.39) in (a,c) such that $\eta(a) = \eta(c) = 0$.*

Theorem 1.3.4 (Jacobi-1830) *Let $L = L(t,x,q)$ be a $C^3([a,b] \times \mathbb{R}^n \times \mathbb{R}^n)$ Lagrangian and let $\bar{\xi}$ be a C^2 extremal of problem (1.14). Then*

I) if $P(t) := L_{qq}(t,\bar{\xi}(t),\bar{\xi}'(t)) > 0$, $t \in [a,b]$, and if no conjugate points exist in $(a,b]$, then $\bar{\xi}$ is a local minimizer.

II) if $\bar{\xi}$ is a local minimizer and $P(t) > 0$ for any $t \in [a,b]$, then there are no conjugate points in (a,b).

Let us apply Theorem 1.3.4 in the minimization problem given in Example 1.3.2.

Example 1.3.5 Consider again the problem

$$\min \left\{ \int_0^T (\xi'(t)^2 - \xi(t)^2)\ dt \mid \xi \in C^1([0,T]),\ \xi(0) = \xi(T) = 0 \right\}.$$

Let us find the values of T for which $\bar{\xi} \equiv 0$ is a local minimizer.
The Euler equation for such a problem is: $\xi''(t) + \xi(t) = 0$, with boundary conditions

[3] The term almost means that the inequality $g''(0) > 0$ for any $\eta \neq 0$ is not sufficient to conclude that $\bar{\xi}$ is minimizer, because it is not uniform in η. We should work a little more to finish the proof.

$\xi(0) = \xi(T) = 0$. Hence, any extremal is of the form $\xi(t) = \alpha \sin(t)$, with $\alpha \sin(T) = 0$. Now, if $T < \pi$, the unique extremal is the zero function, since the latter condition gives $\alpha = 0$. Therefore, $\xi \equiv 0$ is the unique possible (global) minimizer. So, let us consider the Jacobi equation for this problem. It is easy to see that such an equation coincides (because of the special structure of L) with the Euler one and that no conjugate points exist for $T < \pi$. Hence, $\bar{\xi} \equiv 0$ is the unique minimizer of the given problem.
If $T = \pi$, then for any $\alpha \in \mathbb{R}$ the function $\xi_\alpha(t) = \alpha \sin(t)$ is an extremal and a nontrivial solution of the Jacobi equation. Moreover, since

$$J(\xi_\alpha) = \int_0^\pi \alpha^2(\cos^2(t) - \sin^2(t))\ dt = \int_0^\pi \alpha^2 \cos(2t)\ dt = 0,$$

either all the extremals are local minimizers or none of them is. In this case Tonelli's Theorem and Jacobi's result do not apply and so, in particular, we cannot conclude if the 0 function is a minimizer or not. However, by means of Fourier series, it can be shown that for any $\xi \in C^1_*([0,\pi])$ such that $\xi(0) = \xi(\pi) = 0$

$$\int_0^\pi \xi(s)^2\ ds \leq \int_0^\pi \xi'(s)^2\ ds,$$

yielding that any extremal ξ_α is a global minimizer. Indeed, given any $\xi \in C^1_*([0,\pi])$ such that $\xi(0) = \xi(\pi) = 0$, we can define an odd and 2π-periodic function $\tilde{\xi} : [-\pi, \pi] \to \mathbb{R}$ by setting $\tilde{\xi}(t) = \xi(t)$ on $[0, \pi]$ and $\tilde{\xi}(t) = -\xi(-t)$ on $[-\pi, 0)$. For such a function, the Fourier series exists, it is uniformly convergent on $[-\pi, \pi]$ and, due to the fact that we are dealing with an odd function, it is given by

$$\sum_{k \geq 0} b_k \sin(kt), \qquad t \in [-\pi, \pi].$$

Hence,

$$\int_0^\pi \xi(t)^2\ dt = \sum_{k\geq 0} \int_0^\pi b_k^2 \sin^2(kt)\ dt = \frac{\pi}{2} \sum_{k \geq 0} b_k^2$$

and

$$\int_0^\pi \xi'(t)^2\ dt = \sum_{k\geq 0} \int_0^\pi k^2 b_k^2 \cos^2(kt)\ dt = \frac{\pi}{2} \sum_{k \geq 0} k^2 b_k^2.$$

We conclude that for any $\xi \in C^1_*([0,\pi])$ such that $\xi(0) = \xi(\pi) = 0$,

$$\int_0^\pi \xi(s)^2\ ds \leq \int_0^\pi \xi'(s)^2\ ds.$$

It only remains to analyse the case $T > \pi$. To this end, let us consider as above the Jacobi equation for this problem. It is easy to see that any point $c = k\pi$, for $k \in \mathbb{N}$ such that $k\pi < T$ is a conjugate point to 0. By Theorem 1.3.4 we can conclude that $\bar{\xi} \equiv 0$ cannot be a local minimizer for J.

Let us go back to problem (1.14) and to Jacobi's statement. Let $L = L(t, x, q)$ be a $C^3([a, b] \times \mathbb{R}^n \times \mathbb{R}^n)$ Lagrangian and $\bar{\xi}$ be a local minimizer of problem (1.14) for which

$$P(t) := L_{qq}(t, \bar{\xi}(t), \bar{\xi}'(t)) > 0, \qquad t \in [a, b].$$

Hence for any $\eta \in C^1_*([a,b],\mathbb{R}^n)$ such that $\eta(a) = \eta(b) = 0$ we have (actually $\bar{\xi}$ is C^2)

$$0 \leq \frac{d^2}{d\lambda^2} J(\bar{\xi} + \lambda\eta)|_{\lambda=0} = \int_a^b [L_{xx}(t,\bar{\xi}(t),\bar{\xi}'(t))\eta(t)\cdot\eta(t) + L_{qq}(t,\bar{\xi}(t),\bar{\xi}'(t))\eta'(t)\cdot\eta'(t)$$

$$+ 2L_{xq}(t,\bar{\xi}(t),\bar{\xi}'(t))\eta'(t)\cdot\eta(t)]\ dt.$$

Now define

$$\mathcal{L}^{\bar{\xi}}(t,y,z) = L_{xx}(t,\bar{\xi}(t),\bar{\xi}'(t))y\cdot y + L_{qq}(t,\bar{\xi}(t),\bar{\xi}'(t))z\cdot z + 2L_{xq}(t,\bar{\xi}(t),\bar{\xi}'(t))z\cdot y,$$

which is called the *accessory Lagrangian* to problem (1.14). Then consider the new problem

$$\min\left\{\int_a^b \mathcal{L}^{\bar{\xi}}(t,\eta(t),\eta'(t))\ dt \mid \eta \in C^1_*([a,b];\mathbb{R}^n),\ \eta(a) = \eta(b) = 0\right\} \tag{1.40}$$

(here $J_{\bar{\xi}}[\eta] = \int_a^b \mathcal{L}^{\bar{\xi}}(t,\eta(t),\eta'(t))\ dt$ is called the *accessory functional*).
Observe first that the zero function is a global minimizer for this problem and that $J_{\bar{\xi}}[0] = 0$. Moreover, since $\frac{\partial^2\mathcal{L}^{\bar{\xi}}}{\partial z^2}(t,\eta(t),z) = \frac{\partial^2 L}{\partial q^2}(t,\bar{\xi}(t),\bar{\xi}'(t)) > 0$ for all $z \in \mathbb{R}^n$, any other extremal η of problem (1.40) (in the sense of Definition 1.2.11) is a C^1 function and satisfies the classical Euler equation by Corollary 1.2.13. Furthermore, the Euler equation for problem (1.40) is

$$\frac{d}{dt}\left\{\bar{L}_{qq}(t)\eta'(t) + \bar{L}_{xq}(t)\eta(t)\right\} = \bar{L}_{xq}(t)\eta'(t) + \bar{L}_{xx}(t)\eta(t).$$

where

$$\bar{L}_{qq}(t) := L_{qq}(t,\bar{\xi}(t),\bar{\xi}'(t)),\ \bar{L}_{xx}(t) := L_{xx}(t,\bar{\xi}(t),\bar{\xi}'(t)) \text{ and } \bar{L}_{xq}(t) := L_{xq}(t,\bar{\xi}(t),\bar{\xi}'(t)).$$

Any solution of the previous equation is called the *accessory extremal.* It is easy to prove the following connection between accessory extremals and conjugate points.

Proposition 1.3.6 *Let $L = L(t,x,q)$ be a $C^3([a,b]\times\mathbb{R}^n\times\mathbb{R}^n)$ Lagrangian and $\bar{\xi}$ be an extremal of problem (1.14) for which $\bar{L}_{qq}(t) := L_{qq}(t,\bar{\xi}(t),\bar{\xi}'(t)) > 0$, $t \in [a,b]$. Then a point $c \in (a,b]$ is conjugate to a if and only if there exists a nontrivial accessory extremal η_c on $[a,c]$ such that $\eta_c(a) = \eta_c(c) = 0$.*

Remark 1.3.7 Nontriviality of the accessory extremal is equivalent to the condition $\eta_c'(a) \neq 0$ ($\eta_c'(c) \neq 0$) since the Euler equation related to the accessory functional is a second order linear equation.

Now we are ready to prove the Jacobi Necessary Condition.

Proof of Theorem 1.3.4-II)– Let $\bar{\xi}$ be a C^2 local minimizer such that $\bar{L}_{qq}(t) > 0$ and suppose by contradiction that there exists a conjugate point $c \in (a, b)$. Take the nontrivial accessory extremal η_c associated with c as in Proposition 1.3.6. Then,

$$\begin{aligned}
&\int_a^c \mathcal{L}^{\bar{\xi}}(t, \eta_c(t), \eta_c'(t))\, dt \\
= \;& \int_a^c [\bar{L}_{qq}(t)\eta_c'(t) \cdot \eta_c'(t) + 2\bar{L}_{xq}(t)\eta_c'(t) \cdot \eta_c(t) + \bar{L}_{xx}(t)\eta_c(t) \cdot \eta_c(t)]\, dt \\
= \;& \int_a^c \Big\{ \frac{d}{dt}[(\bar{L}_{qq}(t)\eta_c'(t) + \bar{L}_{xq}(t)\eta_c(t)) \cdot \eta_c(t)] - \eta_c(t) \cdot \frac{d}{dt}(\bar{L}_{qq}(t)\eta_c'(t) + \bar{L}_{xq}(t)\eta_c(t)) \\
& \quad + \bar{L}_{xq}(t)\eta_c'(t) \cdot \eta_c(t) + \bar{L}_{xx}(t)\eta_c(t) \cdot \eta_c(t) \Big\}\, dt \\
= \;& \Big[(\bar{L}_{qq}(t)\eta_c'(t) + \bar{L}_{xq}(t)\eta_c(t)) \cdot \eta_c(t) \Big]_a^c = 0,
\end{aligned}$$

since $\eta_c(a) = \eta_c(c) = 0$ and

$$\frac{d}{dt}\Big\{ \bar{L}_{qq}(t)\eta_c'(t) + \bar{L}_{xq}(t)\eta_c(t) \Big\} = \bar{L}_{xq}(t)\eta_c'(t) + \bar{L}_{xx}(t)\eta_c(t).$$

If we define

$$\tilde{\eta}(t) = \begin{cases} \eta_c(t) & \text{for } t \in [a, c], \\ 0 & \text{for } t \in [c, b], \end{cases}$$

we then have that $\tilde{\eta}(t) \in C^1_*([a, b]; \mathbb{R}^n)$, $\tilde{\eta}(a) = \tilde{\eta}(b) = 0$ and

$$J_{\bar{\xi}}[\tilde{\eta}] = \int_a^c \mathcal{L}^{\bar{\xi}}(t, \eta_c(t), \eta_c'(t))\, dt + \int_c^b \mathcal{L}^{\bar{\xi}}(t, 0, 0)\, dt = 0,$$

that is $\tilde{\eta}$ is a minimizer of $J_{\bar{\xi}}$. By Corollary 1.2.13 we conclude that $\tilde{\eta}$ is a C^1 function, against the fact that

$$0 = \lim_{t \to c^+} \tilde{\eta}'(t) \neq \lim_{t \to c^-} \tilde{\eta}'(t) = \eta_c'(c).$$

(see Remark 1.3.7). ∎

Let us now prove the Jacobi Sufficient Condition in the one dimensional case.

Proof of Theorem 1.3.4-I)– Let $\bar{\xi}$ be a C^2 extremal of problem (1.14) such that $\bar{L}_{qq}(t) := L_{qq}(t, \bar{\xi}(t), \bar{\xi}'(t)) > 0$ for all $t \in [a, b]$ and such that no conjugate points exist in $(a, b]$. Consider any $\eta \in C^1_*([a, b])$ with $\eta(a) = \eta(b) = 0$. Since $\bar{\xi}$ is an extremal, then

$$J[\bar{\xi} + \eta] - J[\bar{\xi}] = \frac{1}{2}\frac{d^2}{d\lambda^2} J[\bar{\xi} + \lambda\eta] \qquad \text{for some } \lambda \in (0, 1).$$

Denoting by $L^\lambda_{qq}(t)$ the derivative $L_{qq}(t, \bar{\xi}(t) + \lambda\eta(t), \bar{\xi}'(t) + \lambda\eta'(t))$ we have

$$J[\bar\xi+\eta]-J[\bar\xi]$$

$$=\frac{1}{2}\int_a^b\left[L^\lambda_{qq}(t)\eta'^2(t)+2L^\lambda_{xq}(t)\eta'(t)\eta(t)+L^\lambda_{xx}(t)\eta^2(t)\right]dt$$

$$=\frac{1}{2}J_{\bar\xi}[\eta]+I_0,$$

where

$$I_0 := \frac{1}{2}\int_a^b\Big[(L^\lambda_{qq}(t)-\bar L_{qq}(t))\eta'^2(t)+2(L^\lambda_{xq}(t)-\bar L_{xq}(t))\eta'(t)\eta(t)$$
$$+(L^\lambda_{xx}(t)-\bar L_{xx}(t))\eta^2(t)\Big]\,dt.$$

Since by assumption L is a C^3 function, there exists a constant $K=K(\bar\xi)>0$ such that

$$\left|(L_{qq}(t,\bar\xi(t)+x,\bar\xi'(t)+q)-L_{qq}(t,\bar\xi(t),\bar\xi'(t))\right|\le K(|x|+|q|)$$

for any $(x,q)\in[-1,1]\times[-1,1]$. Hence,

$$\left|(L^\lambda_{qq}(t)-\bar L_{qq}(t))\eta'^2(t)\right|\le K(\|\eta\|_\infty+\|\eta'\|^*_\infty)\eta'^2(t),$$

for any $\eta\in C^1_*([a,b])$ such that $\eta(a)=\eta(b)=0$, provided $\|\eta\|_\infty<1$ and $\|\eta'\|^*_\infty<1$. Similar estimates hold for all the other terms in I_0, giving (increasing K if needed)

$$|I_0|\le\frac{K}{2}(\|\eta\|_\infty+\|\eta'\|^*_\infty)\int_a^b\Big(\eta'^2(t)+2|\eta'(t)\eta(t)|+\eta^2(t)\Big)\,dt.$$

Moreover, as in Wirtinger's Example (see 1.3.2), we have

$$\int_a^b\Big(\eta'^2(t)+2|\eta'(t)\eta(t)|+\eta^2(t)\Big)\,dt\le C\int_a^b\eta'^2(t)\ dt,$$

for some $C=C(a,b)$. So, for any $\varepsilon>0$ we can find $C_\varepsilon>0$ such that, for any $\eta\in C^1_*([a,b])$ satisfying $\eta(a)=\eta(b)=0$ and $\|\eta\|_\infty+\|\eta'\|^*_\infty<C_\varepsilon$, we have

$$|I_0|\le\frac{\varepsilon}{2}\int_a^b\eta'^2(t)\ dt.$$

Then

$$J[\bar\xi+\eta]-J[\bar\xi]\ge\frac{1}{2}J_{\bar\xi}[\eta]-\frac{\varepsilon}{2}\int_a^b\eta'^2(t)\ dt$$

$$=\frac{1}{2}\int_a^b\left[(\bar L_{qq}(t)-\varepsilon)\eta'^2(t)+2\bar L_{xq}(t)\eta'(t)\eta(t)+\bar L_{xx}(t)\eta^2(t)\right]dt$$

$$=\frac{1}{2}\int_a^b\left[(P(t)-\varepsilon)\eta'^2(t)+Q(t)\eta^2(t)\right]dt$$

(where $Q(t) := L_{xx}(t, \bar{\xi}(t), \bar{\xi}'(t)) - \frac{d}{dt}\{L_{xq}(t, \bar{\xi}(t), \bar{\xi}'(t))\}$ and $P(t) := L_{qq}(t, \bar{\xi}(t), \bar{\xi}'(t))$ as in the false Theorem of Legendre). In order to complete the proof, we need the following lemma.

Lemma 1.3.8 *Assume the hypotheses of Theorem 1.3.4-I). Then there exists $\varepsilon_0 > 0$ such that for any $0 < \varepsilon < \varepsilon_0$ the equation $\{(P(t) - \varepsilon)u'(t)\}' + Q(t)u(t) = 0$ admits a (classical) solution $u_\varepsilon \neq 0$ for all $t \in [a, b]$.*

Before proving Lemma 1.3.8, let us conclude Jacobi's Theorem.

Since $\int_a^b \{w(t)\eta^2(t)\}' dt = 0$ for any continuous w and any $\eta \in C^1_*([a, b])$ such that $\eta(a) = \eta(b) = 0$, taking

$$w(t) = -\frac{(P(t) - \varepsilon)u'_\varepsilon(t)}{u_\varepsilon(t)},$$

with u_ε as in Lemma 1.3.8, we have

$$J[\bar{\xi} + \eta] - J[\bar{\xi}] \geq \frac{1}{2}\int_a^b \Big[(P(t) - \varepsilon)\eta'^2(t) + Q(t)\eta^2(t) + \{w(t)\eta^2(t)\}'\Big]\, dt$$

$$= \frac{1}{2}\int_a^b \Big[(P(t) - \varepsilon)\eta'^2(t) + Q(t)\eta^2(t) + \frac{(-(P(t) - \varepsilon)u'_\varepsilon(t))'}{u_\varepsilon(t)}\eta^2(t)$$

$$+ \frac{(P(t) - \varepsilon)u'^2_\varepsilon(t)}{u^2_\varepsilon(t)}\eta^2(t) - 2\frac{(P(t) - \varepsilon)u'_\varepsilon(t)}{u_\varepsilon(t)}\eta(t)\eta'(t)\Big] dt$$

$$= \frac{1}{2}\int_a^b (P(t) - \varepsilon)\Big[\eta'^2(t) + \frac{u'^2_\varepsilon(t)}{u^2_\varepsilon(t)}\eta^2(t) - 2\frac{u'_\varepsilon(t)}{u_\varepsilon(t)}\eta(t)\eta'(t)\Big]\, dt$$

$$= \frac{1}{2}\int_a^b (P(t) - \varepsilon)\Big[\eta'(t) - \frac{u'_\varepsilon(t)}{u_\varepsilon(t)}\eta(t)\Big]^2\, dt \geq 0,$$

for any ε sufficiently small. Hence we have proved that for any $\varepsilon > 0$ small enough, there exists $C_\varepsilon > 0$ such that for any $\eta \in C^1_*([a, b])$ with $\eta(a) = \eta(b) = 0$ and $\|\eta\|_\infty + \|\eta'\|^*_\infty < C_\varepsilon$ we have

$$J[\bar{\xi} + \eta] \geq J[\bar{\xi}],$$

i.e. $\bar{\xi}$ is a local minimizer for functional J. ∎

Proof of Lemma 1.3.8– Let u_0 be the unique solution of the Cauchy problem

$$\begin{cases} -\{P(t)u'(t)\}' + Q(t)u(t) = 0 \quad t \in (a, b) \\ \\ u(a) = 0 \\ \\ u'(a) = 1. \end{cases}$$

Since there are no conjugate points (to a) in $(a, b]$ and $u_0'(a) \neq 0$, then $u_0(t) \neq 0$ for any $t \neq a$. Hence there exist $\delta > 0$ and $d \in (a, b)$ such that $u_0'(t) > \delta$ on $[a, d]$ and $u_0(t) > \delta$ on $[d, b]$. For any $\alpha > 0$ and $\varepsilon > 0$ let $u_{\alpha,\varepsilon}$ be the solution of the perturbed Cauchy problem

$$\begin{cases} -\{(P(t) - \varepsilon)u'(t)\}' + Q(t)u(t) = 0 \quad t \in (a, b) \\ \\ u(a) = \alpha \\ \\ u'(a) = 1. \end{cases}$$

By the continuous dependence of solutions of Cauchy problems from initial data and parameters, we have that for any α, ε sufficiently small,

$$|u_{\alpha,\varepsilon}(t) - u_0(t)| \leq \frac{\delta}{2}, \quad |u'_{\alpha,\varepsilon}(t) - u_0'(t)| \leq \frac{\delta}{2} \qquad \text{for any } t \in [a, b]$$

and then

$$u_{\alpha,\varepsilon}(t) > \frac{\delta}{2}, \qquad \text{on } [d, b],$$

and

$$u'_{\alpha,\varepsilon}(t) > \frac{\delta}{2} \qquad \text{on } [a, d].$$

This obviously implies that $u_{\alpha,\varepsilon}(t) > 0$ on $[a, b]$ for all α, ε sufficiently small. ∎

At the end of this section we consider again the Evans model presented in Example 1.2.24.

Example 1.3.9 We recall that the problem was to maximize

$$\Pi(P) = \int_0^T \left(-2P^2(t) + 3P(t)P'(t) + 4P(t) - P'^2(t) - 3P'(t) - 3\right) dt$$

subject to

$$P(0) = P_0, \ P(T) = P_T, \ P_0 \neq P_T,$$

and that in Example 1.2.24 we got a unique extremal (see (1.32)). We could not conclude that such an extremal was a local maximizer of the problem, because assumption (E1) of Tonelli's Theorem was not fulfilled. However, a positive answer is given by Jacobi's Theorem 1.3.4, since L is strictly concave with respect to q and no conjugate points to 0 exist in the interval $(0, T]$. Indeed, for any fixed $c \in (0, T]$, Jacobi's equation is

$$\eta''(t) - 2\eta(t) = 0,$$

with boundary conditions $\eta(0) = \eta(c) = 0$. It is easy to see that the null function is the unique solution to the previous problem for any c.

Exercise 1.3.10 *Consider again the minimization problem of Exercise 1.2.34. The functional J and the boundary conditions are:*

$$J(\xi) = \int_0^1 (1 + t)(\xi'(t)^2 + \xi(t)) \, dt \qquad \xi(0) = 0, \, \xi(1) = 1,$$

where J is defined on the class of C^1 functions $\xi : [0, 1] \to \mathbb{R}$. By means of Jacobi's Theorem 1.3.4, determine the nature of the extremals.

1.4 Isoperimetric Problem

There are many problems in Calculus of Variations which cannot be formulated as the ones we have introduced so far. Most of them are optimization problems under constraints. The isoperimetric problem is one of these. It can be stated as follows: minimize the functional

$$J[\xi] = \int_a^b L(t, \xi(t), \xi'(t))\ dt$$

subject to

$$\xi \in C^1_*([a,b], \mathbb{R}^n), \qquad \xi(a) = A,\ \xi(b) = B,$$

and

$$G(\xi) = \int_a^b g(t, \xi(t), \xi'(t))\ dt = C, \qquad C \in \mathbb{R}.$$

For such a problem we have the following sufficient condition for optimality:

Theorem 1.4.1 *Let $L, g \in C^1([a,b] \times \mathbb{R}^n \times \mathbb{R}^n)$. Moreover, suppose that for some $\lambda \in \mathbb{R}$ the augmented Lagrangian $\overline{L}$:*

$$\overline{L}(t, x, q) = L(t, x, q) + \lambda g(t, x, q),$$

is a convex function with respect to the (x, q) variable for any fixed $t \in [a, b]$. Then any extremal $\bar{\xi}$ of $\overline{L}$ such that $\bar{\xi}(a) = A$, $\bar{\xi}(b) = B$ is a solution of the isoperimetric problem for

$$C = C_\lambda = \int_a^b g(t, \bar{\xi}(t), \bar{\xi}'(t))\ dt.$$

Proof– Let $\overline{J}[\xi] = J[\xi] + \lambda G(\xi)$. By assumption, $\overline{J}$ is a convex and smooth functional on $C^1_*([a,b], \mathbb{R}^n)$. Furthermore, if $\bar{\xi}$ is an extremal for the augmented Lagrangian such that $\bar{\xi}(a) = A$, $\bar{\xi}(b) = B$, then $\bar{\xi}$ is a minimizer of $\overline{J}$, that is

$$\overline{J}[\bar{\xi}] \le \overline{J}[\xi], \qquad \text{for all} \quad \xi \in C^1_*([a,b], \mathbb{R}^n), \quad \xi(a) = A,\ \xi(b) = B.$$

Indeed, by the convexity and regularity of $\overline{L}$ we have

$$\begin{aligned}
\overline{J}[\xi] - \overline{J}[\bar{\xi}] &= \int_a^b \{\overline{L}(t, \xi(t), \xi'(t)) - \overline{L}(t, \bar{\xi}(t), \bar{\xi}'(t))\}\ dt \\
&\ge \int_a^b \{\overline{L}_x(t, \bar{\xi}(t), \bar{\xi}'(t)) \cdot (\xi(t) - \bar{\xi}(t))\}\ dt \\
&+ \int_a^b \{\overline{L}_q(t, \bar{\xi}(t), \bar{\xi}'(t)) \cdot (\xi'(t) - \bar{\xi}'(t))\}\ dt \\
&= \Big[\overline{L}_q(t, \bar{\xi}(t), \bar{\xi}'(t)) \cdot (\xi(t) - \bar{\xi}(t))\Big]_a^b \\
&+ \int_a^b \left\{-\frac{d}{dt}\overline{L}_q(t, \bar{\xi}(t), \bar{\xi}'(t)) + \overline{L}_x(t, \bar{\xi}(t), \bar{\xi}'(t))\right\} \cdot (\xi(t) - \bar{\xi}(t))\ dt = 0.
\end{aligned}$$

We have just proved that

$$J[\bar{\xi}] + \lambda G(\bar{\xi}) \le J[\xi] + \lambda G(\xi), \qquad \text{for all} \quad \xi \in C^1_*([a,b], \mathbb{R}^n), \quad \xi(a) = A, \ \xi(b) = B.$$

Restricting the previous inequality to the functions $\xi \in C^1_*([a,b], \mathbb{R}^n)$ such that $\xi(a) = A$, $\xi(b) = B$ and $G(\xi) = G(\bar{\xi})$, we get the desired result. ■

Remark 1.4.2 The isoperimetric problem can be also formulated with a finite number of conditions, say

$$G_i(\xi) = \int_a^b g_i(t, \xi(t), \xi'(t))\ dt = C_i, \qquad C_i \in \mathbb{R}, \ i = 1, \ldots, N.$$

In this case, Theorem 1.4.1 holds true provided we consider N scalars λ_i and we define the augmented Lagrangian by $\overline{L} = L + \sum_{i=1}^N \lambda_i g_i$.

We stress that all the regularity results of Section 1.2.2 can be applied to the augmented Lagrangian whenever it satisfies the hypotheses of injectivity and regularity. In the sequel we will present several examples of isoperimetric problems. Anytime that the regularity results apply, we will automatically set the problem on the class of C^1 functions.

Example 1.4.3 Consider the problem

$$\min \left\{ \int_0^1 \xi'(t)^2\ dt \mid \xi : [0,1] \to \mathbb{R}, \xi(0) = \xi(1) = 0, \int_0^1 \xi(t)\ dt = 1 \right\}.$$

Then, $L(q) = q^2$, $\lambda g(x) = \lambda x$ are convex functions for any $\lambda \in \mathbb{R}$. The Euler equation for the augmented Lagrangian is $2x''(t) = \lambda$ ($\overline{L}$ is strictly convex with respect to q and regular). Hence, for λ fixed, there is a unique extremal of the augmented Lagrangian which satisfies the boundary conditions, namely $\bar{\xi}(t) = \frac{\lambda}{4}(t^2 - t)$. Since

$$\int_0^1 \bar{\xi}(t)\ dt = \left[\frac{\lambda}{4} \left(\frac{t^3}{3} - \frac{t^2}{2} \right) \right]_0^1 = -\frac{\lambda}{24},$$

for $\lambda = -24$ we obtain that $\bar{\xi}(t) = 6(t - t^2)$ is a minimizer for the original problem.

Example 1.4.4 (Queen Dido's Problem) It deals with the maximization of the functional

$$\bar{J}[\xi] = \int_{-a}^a \xi(t)\ dt, \qquad (a > 0 \text{ fixed})$$

subject to

$$\xi(-a) = \xi(a) = 0 \text{ and } \int_{-a}^a \sqrt{1 + \xi'(t)^2}\ dt = L,$$

where $2a < L < \pi a$. In other words, we want to maximize the area of the region enclosed by the curve $\{(t, \xi(t)) \ : \ t \in [-a, a]\}$ and the line segment $[-a, a]$ among all the curves with fixed extremes $(-a, \xi(-a)) = (-a, 0)$, $(a, \xi(a)) = (a, 0)$ and prescribed length L.

Setting $J[\xi] := -\bar{J}[\xi]$, the hypotheses of Theorem 1.4.1 are fulfilled provided $\lambda \geq 0$. Let $\bar{\xi}$ be an extremal for the augmented Lagrangian

$$\overline{L}(t,x,q) = -x + \lambda\sqrt{1+q^2},$$

which is a regular function and it is strictly convex with respect to q. The Euler equation is

$$\frac{d}{dt}\left[\frac{\lambda\bar{\xi}'(t)}{\sqrt{1+\bar{\xi}'(t)^2}}\right] = -1.$$

Integrating we obtain

$$\frac{\lambda\bar{\xi}'(t)}{\sqrt{1+\bar{\xi}'(t)^2}} = c - t,$$

and squaring the above equality we get

$$\bar{\xi}'(t)^2(\lambda^2 - (c-t)^2) = (c-t)^2.$$

Simple maximality arguments imply that we can restrict to symmetric solutions, decreasing for $t > 0$; hence $c = 0$ and

$$\bar{\xi}'(t) = \frac{-t}{\sqrt{\lambda^2 - t^2}}, \qquad t \in [-a, a]$$

provided $\lambda > a$. Integrating again and imposing $\bar{\xi}(a) = 0$ we get

$$\bar{\xi}(t) = \sqrt{\lambda^2 - t^2} - \sqrt{\lambda^2 - a^2}.$$

Hence, for any $\lambda > a$, the solution is a circle whose equation is

$$(\bar{\xi}(t) + \sqrt{\lambda^2 - a^2})^2 + t^2 = \lambda^2.$$

In order to find the minimizers of Dido's problem, we have to find (if they exist) the values of λ such that $\int_{-a}^{a}\sqrt{1+\bar{\xi}'(t)^2}\,dt = L$. This means that λ must satisfy

$$\begin{aligned} L &= \int_{-a}^{a}\sqrt{1+\frac{t^2}{\lambda^2 - t^2}}\,dt \\ &= \lambda\int_{-a}^{a}\frac{dt}{\sqrt{\lambda^2 - t^2}} \qquad (\lambda > a) \\ &= 2\lambda\arcsin\frac{a}{\lambda}. \end{aligned}$$

First notice that equation $L = 2\lambda\arcsin\frac{a}{\lambda}$ is well-posed, since $\lambda > a$ and $L < \pi a$. Let us now consider the map

$$f(\lambda) = \frac{\sin\frac{L}{2\lambda}}{\frac{L}{2\lambda}}\frac{L}{2} \qquad \lambda > a.$$

Since $\lim_{\lambda\to+\infty} f(\lambda) = \frac{L}{2} > a$ and $\lim_{\lambda\to a^+} f(\lambda) = a\sin\frac{L}{2a} < a$, we conclude the existence of $\lambda > a$ such that $f(\lambda) = a$, that is $L = 2\lambda\arcsin\frac{a}{\lambda}$.

The following theorem is a necessary condition for the local minimizers of the isoperimetric problem in the case when no convexity assumptions are made on the data L and g.

Theorem 1.4.5 *Let $L, g \in C^1([a,b]\times\mathbb{R}^n\times\mathbb{R}^n)$. If $\bar{\xi}$ is a local solution of the isoperimetric problem, then either*

(i) $\bar{\xi}$ is an extremal for the functional G, i.e. $\frac{d}{ds}G(\bar{\xi}+s\eta)|_{s=0} = 0$ for any $\eta \in C^1_*([a,b],\mathbb{R}^n)$ such that $\eta(a)=\eta(b)=0$;

or

(ii) there exists $\lambda \in \mathbb{R}$ such that $\bar{\xi}$ is an extremal for $J+\lambda G$.

Proof– Let us set $\mathcal{V} := \{\eta \in C^1_*([a,b],\mathbb{R}^n) \ : \ \eta(a)=\eta(b)=0\}$. We will first prove that if $\bar{\xi}$ is a local solution of the isoperimetric problem, then for any $\eta, \mu \in \mathcal{V}$

$$\det \begin{pmatrix} J'(\bar{\xi})\eta & J'(\bar{\xi})\mu \\ \\ G'(\bar{\xi})\eta & G'(\bar{\xi})\mu \end{pmatrix} = 0,$$

where $J'(\bar{\xi})\eta := \frac{d}{ds}J(\bar{\xi}+s\eta)|_{s=0}$. Indeed, suppose by contradiction that there exist $\eta, \mu \in \mathcal{V}$ such that

$$\det \begin{pmatrix} J'(\bar{\xi})\eta & J'(\bar{\xi})\mu \\ \\ G'(\bar{\xi})\eta & G'(\bar{\xi})\mu \end{pmatrix} \neq 0$$

and for $r>0$ small consider the map $\varphi : B_r(0,0) \in \mathbb{R}^2 \to \mathbb{R}^2$, defined by

$$\varphi(l,m) = (J(\bar{\xi}+l\eta+m\mu), G(\bar{\xi}+l\eta+m\mu)).$$

Then, $\varphi(0,0) = (J(\bar{\xi}), G(\bar{\xi}))$, $\varphi \in C^1(B_r(0,0))$ because of the regularity of L and g, and $\det D\varphi(0,0) \neq 0$ by the previous assumption. Hence, for a suitable $r' \leq r$, φ is a local diffeomorphism between $B_{r'}(0,0)$ and $\varphi(B_{r'}(0,0))$ and then it is possible to find some $(l,m) \in B_{r'}(0,0)$ such that $G(\bar{\xi}+l\eta+m\mu) = G(\bar{\xi})$ and $J(\bar{\xi}+l\eta+m\mu) < J(\bar{\xi})$, against the fact that $\bar{\xi}$ is a local solution of the isoperimetric problem.
Let us now conclude the proof of the theorem. If $\bar{\xi}$ is not an extremal for G, then there exists $\bar{\mu} \in \mathcal{V}$ such that $G'(\bar{\xi})\bar{\mu} \neq 0$. Since for any $\eta \in \mathcal{V}$ we have

$$\det \begin{pmatrix} J'(\bar{\xi})\eta & J'(\bar{\xi})\bar{\mu} \\ \\ G'(\bar{\xi})\eta & G'(\bar{\xi})\bar{\mu} \end{pmatrix} = 0,$$

then $J'(\bar{\xi})\eta G'(\bar{\xi})\bar{\mu} = J'(\bar{\xi})\bar{\mu}G'(\bar{\xi})\eta$, i.e.

$$J'(\bar{\xi})\eta = \frac{J'(\bar{\xi})\bar{\mu}}{G'(\bar{\xi})\bar{\mu}}G'(\bar{\xi})\eta = -\lambda G'(\bar{\xi})\eta, \quad \forall \eta \in \mathcal{V}, \qquad \text{if we set } \lambda := -\frac{J'(\bar{\xi})\bar{\mu}}{G'(\bar{\xi})\bar{\mu}},$$

which is (ii). ∎

Example 1.4.6 In Example 1.4.4 we needed the assumption $L < a\pi$ in order to use a cartesian representation for the boundary of the region whose area was to be maximized. If we get rid of it, we can formulate Queen Dido's problem in the following way: let $\xi : [0,1] \to \mathbb{R}^2$, $\xi(t) = (\xi_1(t), \xi_2(t))$, be any piecewise regular curve in the plane such that $\xi(0) = (-a, 0)$, $\xi(1) = (a, 0)$. Since by Green's formula the area of the region enclosed by the curve and the line segment $[-a, a]$ is given by $\int_0^1 \xi_1(t)\xi_2'(t)\, dt$, then Queen Dido's problem becomes

$$\max\left\{\int_0^1 \xi_1(t)\xi_2'(t)\, dt \mid \xi : [0,1] \to \mathbb{R}^2,\ \begin{matrix}\xi(0) = (-a,0) \\ \xi(1) = (a,0)\end{matrix}\ , \int_0^1 \sqrt{\xi_1'(t)^2 + \xi_2'(t)^2}\, dt = L\right\}.$$

Now, the Lagrangian $L(x_1, x_2, q_1, q_2) = x_1 q_2$ is not convex, nor injective, so that Theorems 1.4.1 and 1.2.13 do not apply. On the other hand, $g(q_1, q_2) = \sqrt{q_1^2 + q_2^2}$ is not C^1 in a neighbourhood of $(0, 0)$, so that we cannot use Theorem 1.4.5 if we consider the problem on the class of C^1_* functions. However, if we restrict our analysis to the class of the regular functions with nonvanishing gradient, we can still apply Theorem 1.4.5 to find some necessary conditions for maximizers belonging to such a class. Indeed, if $\bar{\xi}$ is a maximizer, then $\bar{\xi}$ is an extremal for either functional G or functional $J + \lambda G$, for some $\lambda \in \mathbb{R}$. In the former case, $\bar{\xi}$ must satisfy equation

$$\frac{d}{dt}\left\{\frac{\bar{\xi}'(t)}{\|\bar{\xi}'(t)\|}\right\} = 0,$$

that is

$$\frac{\bar{\xi}'(t)}{\|\bar{\xi}'(t)\|} = k \qquad \text{for some} \qquad k = (k_1, k_2) \in \mathbb{R}^2,\ \|k\| = 1.$$

Then,

$$\|\bar{\xi}'(t)\| = \langle \bar{\xi}'(t), k\rangle$$

and

$$\int_0^1 \sqrt{\xi_1'(t)^2 + \xi_2'(t)^2}\, dt = \int_0^1 \|\bar{\xi}'(t)\|\, dt = 2ak_1 \le 2a < L,$$

which is a contradiction. In the latter case, the Euler equation is

$$\begin{cases} \dfrac{d}{dt}\left\{\dfrac{\lambda\bar{\xi}_1'(t)}{\|\bar{\xi}'(t)\|}\right\} = -\bar{\xi}_2'(t) \\[2ex] \dfrac{d}{dt}\left\{-\bar{\xi}_1(t) + \dfrac{\lambda\bar{\xi}_2'(t)}{\|\bar{\xi}'(t)\|}\right\} = 0. \end{cases}$$

Hence, there exist $c_1, c_2 \in \mathbb{R}$ such that $\bar{\xi}$ is a solution of the system

$$\begin{cases} \dfrac{\lambda\bar{\xi}_1'(t)}{\|\bar{\xi}'(t)\|} = -\bar{\xi}_2(t) + c_1 \\[2ex] -\bar{\xi}_1(t) + \dfrac{\lambda\bar{\xi}_2'(t)}{\|\bar{\xi}'(t)\|} = c_2. \end{cases}$$

Since $\bar{\xi}$ is a regular curve, then $\xi_1'(t)^2 + \xi_2'(t)^2 \neq 0$ for all $t \in [0, 1]$ and

$$\frac{\lambda \bar{\xi}_1'(t)}{c_1 - \bar{\xi}_2(t)} = \frac{\lambda \bar{\xi}_2'(t)}{c_2 + \bar{\xi}_1(t)}.$$

Now $\lambda \neq 0$, since otherwise $\bar{\xi} \equiv 0$, and hence we obtain $\bar{\xi}_2'(t)(c_1 - \bar{\xi}_2(t)) = \bar{\xi}_1'(t)(c_2 + \bar{\xi}_1(t))$, i.e.

$$\frac{1}{2}\frac{d}{dt}\{\xi_1'(t)^2 + \xi_2'(t)^2 + 2c_2\bar{\xi}_1(t) - 2c_1\bar{\xi}_2(t)\} = 0.$$

This means that there exists some $R > 0$ such that

$$(\bar{\xi}_1(t) + c_2)^2 + (\bar{\xi}_2(t) - c_1)^2 = R - (c_1^2 + c_2^2).$$

We can conclude that if Queen Dido's problem admits a solution $\bar{\xi}$ in the class of regular curves, then it has to be the parameterization of a circular sector.

Exercise 1.4.7 *Using conditions* $\xi(0) = (-a, 0)$, $\xi(1) = (a, 0)$ *and*

$$\int_0^1 \sqrt{\xi_1'(t)^2 + \xi_2'(t)^2}\ dt = L,$$

find out the admissible values for c_1, c_2 *and* R *above.*

An economic example of isoperimetric problem is the Hotelling model to determine the optimal rate of extraction of an exhaustible resource. See also [17].

Example 1.4.8 The Hotelling Model describes the problem of minimizing over the time interval $[0, T]$, the discounted price of extraction of an exhaustible resource when there are B units of resource to extract, that is to minimize

$$\int_0^T e^{-rt} x^2(t)\ dt \tag{1.41}$$

subject to

$$\int_0^T x(t) dt = B, \tag{1.42}$$

where $x(\cdot)$ is the rate of extraction, $r > 0$ and e^{-rt} is the discount factor. The problem can be studied in two ways: either introducing a new state variable or applying the augmented Lagrangian method. We begin with the first approach.
We define a new state variable $y(\cdot)$ as

$$y(t) = \int_0^t x(s)\ ds,$$

so that $y'(t) = x(t)$, $y(0) = 0$ and $y(T) = B$.
Then problem (1.41)–(1.42) can be reformulated in the following way: to minimize

$$\int_0^T e^{-rt} y'^2(t)\ dt \tag{1.43}$$

subject to $y(0) = 0$ and $y(T) = B$.
Since $L(t,q) = e^{-rt}q^2$ is a regular function and it is strictly convex with respect to q, we have $\frac{\partial L}{\partial x}(t,q) = 0$, $\frac{\partial L}{\partial q}(t,q) = 2qe^{-rt}$ and the Euler equation is

$$\frac{d}{dt}\frac{\partial L}{\partial q}(t, y'(t)) = 2e^{-rt}y''(t) - 2re^{-rt}y'(t) = 0.$$

Solving the second order homogeneous linear differential equation

$$y''(t) - ry'(t) = 0,$$

we get $y(t) = c_1 + c_2e^{rt}$, where c_1 and c_2 are constants to be determined exploiting the initial and the terminal conditions, i.e. solving the linear system

$$\begin{cases} c_1 + c_2 = y(0) = 0 \\ \\ c_1 + e^{rt}c_2 = y(T) = B. \end{cases}$$

Its solution is

$$c_1 = -\frac{B}{e^{rT} - 1} \qquad \text{and} \qquad c_2 = \frac{B}{e^{rT} - 1};$$

hence

$$y(t) = \frac{B}{e^{rT} - 1}[e^{rt} - 1] \qquad \text{and then} \qquad x(t) = y'(t) = \frac{rB}{e^{rT} - 1}e^{rt}.$$

Applying the augmented Lagrangian method, we have $\overline{L}(t,x) = e^{-rt}x^2 + \lambda x$. Therefore $\frac{\partial \overline{L}}{\partial x}(t,x) = 2xe^{-rt} + \lambda$ and $\frac{\partial \overline{L}}{\partial q}(t,x) = 0$. The Euler equation is

$$\int_0^t \frac{\partial \overline{L}}{\partial x}(s, x(s))\, ds = \int_0^t 2x(s)e^{-rs}\, ds + \lambda t = c \qquad \text{for some} \quad c \in \mathbb{R};$$

hence, differentiating with respect to t,

$$2x(t)e^{-rt} + \lambda = 0, \qquad \text{that is} \qquad x(t) = -\frac{\lambda}{2}e^{rt}.$$

The multiplier λ can be found exploiting (1.42):

$$B = \int_0^T x(t)\, dt = -\int_0^T \frac{\lambda}{2}e^{rt}\, dt = -\frac{\lambda}{2r}[e^{rT} - 1].$$

Indeed λ must satisfy

$$\lambda = -\frac{2rB}{e^{rT} - 1} \qquad \text{and then} \qquad x(t) = \frac{rB}{e^{rT} - 1}e^{rt}.$$

Exercise 1.4.9 *Minimize the functional*

$$J(\xi) = \int_0^1 \left[\frac{1}{2}\xi'(s)^2 + \frac{1}{2}\xi(s)^2 - \xi(s)\right] ds$$

on the arcs $\xi \in C^1((0,1))$ *subject to*

$$\xi(0) = \xi(1) = 0$$

and

$$\int_0^1 (\xi'(s) - \xi(s))^2 \, ds = \frac{7e - 5}{4(1+e)}.$$

Chapter 2

Optimal Control Theory

2.1 Basic Problem in Optimal Control Theory

One of the basic problems in Calculus of Variations is to minimize the functional

$$\int_0^T L(t, x(t), x'(t))\; dt + \phi_T(x(T)),$$

on $x(\cdot) \in A$, where

$$A = \{x \in C^1_*([0,T]; \mathbb{R}^n) : x(0) = x_0,\ x_0 \in \mathbb{R}^n\}.$$

This problem can be rewritten as

$$\min_{u(\cdot)} \int_0^T L(t, x(t), u(t))\; dt + \phi_T(x(T)), \tag{2.1}$$

where $u : [0,T] \to \mathbb{R}^n$ is called the *control variable* and the *state variable* $x = x(t; x_0, u)$ (also called *trajectory*) is the solution of the controlled system

$$\begin{cases} x'(t) = u(t) \\ \\ x(0) = x_0. \end{cases} \tag{2.2}$$

In general, the controls u that can be applied to govern the system (2.2) are restricted to take values in a closed subset $U \subset \mathbb{R}^n$, called the *control space*. Calculus of Variations Theory can no longer be used if $U \neq \mathbb{R}^n$.

The state equation (2.2) can be considered in a more general form

$$\begin{cases} x'(t) = f(t, x(t), u(t)) \\ \\ x(0) = x_0, \end{cases} \tag{2.3}$$

where $U \subset \mathbb{R}^k$ is a closed set, $f : [0,T] \times \mathbb{R}^n \times U \to \mathbb{R}$ and $u : [0,T] \to U$ is piecewise continuous (in the sequel $u \in C_*([0,T], U)$).

Under standard hypotheses, the state equation (2.3) admits a unique solution for any fixed control variable. Indeed, the following modification of the classical Picard Theorem holds true.

Lemma 2.1.1 *Suppose that $f : [0,T] \times \mathbb{R}^n \times U \to \mathbb{R}$ is a continuous function and that there exists some $K > 0$ such that for any $t \in [0,T]$, $u \in U$ and $x, y \in \mathbb{R}^n$*

$$|f(t,x,u) - f(t,y,u)| \le K|x-y|, \qquad |f(t,x,u)| \le K(1+|x|). \tag{2.4}$$

Then, for any fixed $u \in C_*([0,T],U)$ there exists a unique solution x to (2.3), that is a piecewise C^1 function x ($x \in C^1_*([0,T],\mathbb{R}^n)$ in the sequel) which satisfies (2.3) up to the discontinuity points of the control variable u.

Notice that (2.4) require f to be Lipschitz continuous and sublinear with respect to x, uniformly in t and u.
Now let us define the basic problem in Optimal Control Theory.

Definition 2.1.2 *Suppose that $U \subset \mathbb{R}^k$ is a closed set, $f \in C([0,T] \times \mathbb{R}^n \times U)$ satisfies hypothesis (2.4) and that $L : [0,T] \times \mathbb{R}^n \times U \to \mathbb{R}$ and $\phi : \mathbb{R}^n \to \mathbb{R}$ are continuous and bounded from below functions. Then problem*

$$\inf_{u(\cdot) \in C_*([0,T],U)} \int_0^T L(t, x^{x_0}(t,u), u(t))\ dt + \phi(x^{x_0}(T,u)), \tag{2.5}$$

where $x^{x_0}(\cdot,u) \in C^1_*([0,T],\mathbb{R}^n)$ is a solution of

$$\begin{cases} x'(t) = f(t, x(t), u(t)) \\ \\ x(0) = x_0, \end{cases} \tag{2.6}$$

is called Bolza problem*; if in (2.5) $L = 0$ or $\phi = 0$, then problem (2.5)–(2.6) is called respectively* Mayer problem *and* Lagrange problem*.*

Example 2.1.3 (Boat in stream) We want to model the problem of a boat leaving from the shore to enter in a wide basin where the current is parallel to the shore, increasing with the distance from it. We assume that the engine can move the boat in any direction, but its power is limited. For a fixed time $T > 0$, we want to evaluate the furthest point the boat can reach from the starting point, measured along shore. The mathematical model can be formulated as follows: we fix the (x_1, x_2) Cartesian axes in such a way that the starting point is $(0,0)$ and the x_1–axis coincides with the (starting) shore. Hence, we want to solve the Mayer problem

$$\min\ -x_1(T),$$

where $(x_1, x_2) : [0, T] \to \mathbb{R}^2$ is a solution of the system

$$\begin{cases} x_1'(t) = x_2(t) + u_1(t) \qquad t \in (0, T) \\ x_2'(t) = u_2(t) \qquad t \in (0, T) \\ x_1(0) = 0 \\ x_2(0) = 0, \end{cases}$$

and the controls (u_1, u_2), which represent the engine power, vary in the unit ball of $\mathbb{R}^2$, i.e. $u_1^2 + u_2^2 \leq 1$.

In what follows, we will always assume at least the hypotheses of Definition 2.1.2.

Remark 2.1.4 The Bolza type control problem seems to be the most general formulation. Actually, if we enlarge the dimension of the state space we can rewrite a Bolza problem into a Mayer problem with a higher dimensional state space. In fact, let $X(t) \in \mathbb{R} \times \mathbb{R}^n$ be defined as

$$X(t) = \begin{pmatrix} z(t) \\ x(t) \end{pmatrix},$$

where $z(t) = \int_0^t L(s, x(s), u(s))\, ds$.

Consider the state equation

$$\begin{cases} X'(t) = \begin{pmatrix} z'(t) \\ x'(t) \end{pmatrix} = \begin{pmatrix} L(t, x(t), u(t)) \\ f(t, x(t), u(t)) \end{pmatrix} = F(t, X(t), u(t)); \\ X(0) = \begin{pmatrix} 0 \\ x_0 \end{pmatrix}. \end{cases} \tag{2.7}$$

Then the Bolza problem can be written as

$$\inf_{u(\cdot)} \Phi(X(T)) \tag{2.8}$$

where $\Phi(X(T)) = z(T) + \phi(x(T))$.

On the other hand, a Lagrange problem can be rewritten as a Mayer problem applying the above reasoning and recalling that $\phi = 0$.

Another example of optimal control problems is the one concerning the model of pumping out water from a basin, see [15].

Example 2.1.5 This example is related to the problem of emptying a basin in a fixed amount of time T. The power u of the engine used by the pump is the control variable and the height x of the level of water left in the basin is the state variable.
The functional is the cost to pay to empty the water reservoir. It consists of two terms: the first one is the pumping cost, while the second one is a cost on the residual level of water in the basin.
Let $T > 0$ be fixed and $U = [0, 1]$. The problem is

$$\min_{u(\cdot)} \int_0^T u^2(t)\ dt + x^2(T) \tag{2.9}$$

where $u : [0, T] \to [0, 1]$ and $x = x(t; x_0, u)$ is a solution of the controlled system

$$\begin{cases} x'(t) = -u(t) \\ \\ x(0) = 1. \end{cases} \tag{2.10}$$

2.2 Pontryagin Minimum/Maximum Principle

Let us start by analysing the Mayer problem. Fix $T > 0$, $x_0 \in \mathbb{R}^n$ and consider the minimization problem

$$\inf_{u(\cdot)\in C_*([0,T],U)} \phi(x(T, u)), \tag{2.11}$$

where $x(\cdot, u) \in C^1_*([0, T], \mathbb{R}^n)$ is a solution of

$$\begin{cases} x'(t) = f(t, x(t), u(t)) \\ \\ x(0) = x_0, \end{cases} \tag{2.12}$$

(we omit the dependence of the state variable from x_0 for brevity).

Definition 2.2.1 *Control $\bar{u}$ is said to be* optimal *for problem (2.11)–(2.12) if*

$$\phi_T(x(T, \bar{u})) = \min_{u(\cdot)} \phi_T(x(T, u))$$

and $\bar{x}(t, \bar{u})$ is the optimal trajectory *associated with it. In what follows we will often refer to $\{\bar{u}, \bar{x}\}$ as an* optimal pair*.*

It can be proved (but it is far beyond the scope of these notes) that if the control space U is compact and $f(t, x, U)$ is a convex set for any $t \in [0, T]$ and $x \in \mathbb{R}^n$, then the Mayer problem admits an optimal pair $\{\bar{u}, \bar{x}\}$. But in this section we are interested in determining necessary conditions for optimality. Let us begin with the scalar case. In Calculus of Variations, we considered a perturbation of the optimal trajectory $\bar{\xi}$ and we built a family of admissible arcs. In this case perturbing the optimal trajectory $\bar{x}$ we get $x(t) = \bar{x}(t) + \lambda\eta(t)$, for $\lambda \in \mathbb{R}$ and $\eta : [0, T] \to \mathbb{R}$ and we cannot ensure that the perturbed trajectory x satisfies (2.12); on the other hand, perturbing the control $\bar{u}$ so that $v(t) = \bar{u}(t) + \gamma\delta(t)$, for $\gamma \in \mathbb{R}$ and $\delta : [0,T] \to U$, we cannot be certain that the perturbed control v takes values into U. This is why in Optimal Control Theory we cannot research necessary conditions exploiting the same kind of perturbations as we did in the Calculus of Variations approach. In the present case we will use a particular kind of perturbations (*spike perturbations*), using a constant control on some short interval.

Theorem 2.2.2 *Assume that $\phi \in C^1(\mathbb{R}^n)$ and $D_x f$ is continuous on $[0,T] \times \mathbb{R}^n \times U$. Let $\{\bar{u}, \bar{x}\}$ be an optimal pair for problem (2.11)–(2.12) and let $\lambda(\cdot)$ be the C^1_* solution of*

$$\begin{cases} -\lambda'(t) = (D_x f)^* \left(t, \bar{x}(t,\bar{u}), \bar{u}(t)\right) \cdot \lambda(t) & \text{on } [0,T] \\ \\ \lambda(T) = D\phi(\bar{x}(T,\bar{u})). \end{cases} \tag{2.13}$$

Then,

$$\langle \lambda(t), f(t, \bar{x}(t,\bar{u}), \bar{u}(t))\rangle = \min_{v \in U} \langle \lambda(t), f(t, \bar{x}(t,\bar{u}), v)\rangle, \tag{2.14}$$

in any continuity point t of $\bar{u}$.

The equation (2.14) is called *Pontryagin Minimum Principle*, the equation in (2.13) is called *co–state* or *adjoint equation* and the terminal condition on λ is called *transversality condition*. Function λ is often referred to as *co–state* .
In equation (2.13), the notation $D_x f^*$ stands for the adjoint matrix of $D_x f$.

Proof–For simplicity we will write $\bar{x}(t)$ instead of $\bar{x}(t,\bar{u})$. Fix any continuity point $s \in (0,T]$ of $\bar{u}$, $h \in (0,s)$, $v \in U$, and define

$$u_h(t) = \begin{cases} v & \text{if } s-h < t \le s \\ \\ \bar{u}(t) & \text{elsewhere .} \end{cases}$$

We denote by $x_h(\cdot)$ the corresponding state $x_h(\cdot, u_h)$.
Since $\bar{u}$ is optimal we have

$$\phi(x_h(T)) \ge \phi(\bar{x}(T));$$

hence, dividing by h and passing to the limit as $h \to 0^+$, we get (if the limit exists)

$$0 \le \lim_{h \to 0^+} \frac{\phi(x_h(T) - \phi(\bar{x}(T)}{h} \tag{2.15}$$

In order to prove the existence and to compute the limit in (2.15) we need the following two lemmas.

Lemma 2.2.3 *Assume the hypotheses of Theorem 2.2.2. Under the above assumptions on s and h we have that there exists $C > 0$ such that*

$$|x_h(t) - \bar{x}(t)| \le Ch$$

for all $t \in [0,T]$ and $h \in (0,s)$. In particular, x_h uniformly converges to $\bar{x}$ on $[0,T]$.

Proof– Take $M > 0$ such that $|f(t, \bar{x}(t), \bar{u}(t))| \leq M$ for any $t \in [0, T]$. Then, by the sublinear growth of f we have

$$|x_h(t)| = \left| x_0 + \int_0^t f(r, x_h(r), u_h(r))\, dr \right| \leq |x_0| + \int_0^t K(1 + |x_h(r)|)\, dr,$$

and hence $|x_h(t)| \leq (|x_0| + KT)e^{KT}$ for any $t \in [0, T]$ by the Gronwall inequality[1]. Moreover,

$$x_h(t) - \bar{x}(t) = \int_0^t [f(r, x_h(r), u_h(r)) - f(r, \bar{x}(r), \bar{u}(r))]\, dr = \int_0^t F_h(r)\, dr,$$

where

$$F_h(r) = \begin{cases} 0 & \text{for } r \in (0, s-h) \\ \\ f(r, x_h(r), v) - f(r, \bar{x}(r), \bar{u}(r)) & \text{for } r \in (s-h, s) \\ \\ f(r, x_h(r), \bar{u}(r)) - f(r, \bar{x}(r), \bar{u}(r)) & \text{for } r \in (s, T) \end{cases}$$

Being f and x_h bounded, we easily find that also $|F_h(r)| \leq \tilde{M}$ for all $r \in (s-h, s)$ and some $\tilde{M} > 0$. Hence,

$$|x_h(t) - \bar{x}(t)| \leq \begin{cases} 0 & \text{for } t \in [0, s-h] \\ \\ \tilde{M}h & \text{for } t \in [s-h, s] \\ \\ \tilde{M}h + \displaystyle\int_s^t K|x_h(r) - \bar{x}(r)|\, dr & \text{for } t \in [s, T]. \end{cases}$$

Then, by the Gronwall inequality we have $|x_h(t) - \bar{x}(t)| \leq \tilde{M}he^{K(t-s)}$ on $[s, T]$ and the assertion follows by taking $C = \tilde{M}e^{KT}$. ∎

Lemma 2.2.4 *Assume the hypotheses of Theorem 2.2.2. Under the above assumptions on s and h we have that*

$$\frac{x_h(t) - \bar{x}(t)}{h} \to \xi(t), \qquad \textit{uniformly on } [s, T],$$

where ξ solves

$$\begin{cases} \xi'(t) = D_x f(t, \bar{x}(t), \bar{u}(t)) \cdot \xi(t) & \textit{on } (s, T], \\ \\ \xi(s) = f(s, \bar{x}(s), v) - f(s, \bar{x}(s), \bar{u}(s)). \end{cases}$$

---

[1] **Gronwall's Lemma**
Let $u_0 \in \mathbb{R}$, $K > 0$ and $u : [a, b] \to \mathbb{R}$ be a continuous function. If

$$u(t) \leq u_0 + K \int_a^t u(s)\, ds, \qquad \forall t \in [a, b],$$

then $u(t) \leq u_0 e^{K(t-a)}$ for all $t \in [a, b]$.

Proof– For any $t \in (s, T]$ we have

$$
\begin{aligned}
&\frac{x_h(t)-\bar{x}(t)}{h} - \xi(t) \\
= \; & \frac{1}{h}\int_{s-h}^{s} F_h(r)\,dr + \frac{1}{h}\int_{s}^{t} F_h(r)\,dr - \left(\xi(s) + \int_s^t D_x f(r,\bar{x}(r),\bar{u}(r))\cdot\xi(r)\,dr\right) \\
= \; & \frac{1}{h}\int_{s-h}^{s} [f(r,x_h(r),v) - f(r,\bar{x}(r),\bar{u}(r)) - (f(s,x_h(s),v) - f(s,\bar{x}(s),\bar{u}(s)))]\,dr \\
+ \; & \int_s^t \left[\frac{f(r,x_h(r),\bar{u}(r)) - f(r,\bar{x}(r),\bar{u}(r))}{h} - D_x f(r,\bar{x}(r),\bar{u}(r))\cdot\xi(r)\right] dr \\
=: \; & I_1 + I_2.
\end{aligned}
$$

Now,

$$
\begin{aligned}
I_1 \; = \; & \frac{1}{h}\int_{s-h}^{s} [f(r,x_h(r),v) - f(r,\bar{x}(r),v)]\,dr \\
+ \; & \frac{1}{h}\int_{s-h}^{s} [f(r,\bar{x}(r),v) - f(r,\bar{x}(r),\bar{u}(r))]\,dr - [f(s,x_h(s),v) - f(s,\bar{x}(s),\bar{u}(s))] \\
:= \; & I_{11} + I_{12} + I_{13},
\end{aligned}
$$

where I_{11} is bounded by Ch by Lemma 2.2.3 and $I_{12}+I_{13}$ is of the form $\frac{1}{h}\int_{s-h}^{s} g(r)\,dr - g(s)$ and then goes to 0 as $h \to 0$ because s is a continuity point for $\bar{u}$. Hence we can write $I_1 = \varepsilon(h)$ with $\varepsilon(h) \to 0$ as $h \to 0$. As for I_2 we have

$$
f(r,x_h(r),\bar{u}(r)) - f(r,\bar{x}(r),\bar{u}(r)) = D_x f(r,\bar{x}(r),\bar{u}(r))\cdot(x_h(r)-\bar{x}(r)) + \omega(h),
$$

where $\frac{\omega(h)}{h} \to 0$ as $h \to 0$, because of the regularity of f and the uniform convergence of x_h to $\bar{x}$. Collecting together the previous estimates we deduce that

$$
\left|\frac{x_h(t)-\bar{x}(t)}{h} - \xi(t)\right| \le |\varepsilon(h)| + \int_s^t |D_x f(r,\bar{x}(r),\bar{u}(r))| \left|\frac{x_h(r)-\bar{x}(r)}{h} - \xi(r)\right| dr + T\left|\frac{\omega(h)}{h}\right|.
$$

Using the Gronwall inequality we conclude that

$$
\left|\frac{x_h(t)-\bar{x}(t)}{h} - \xi(t)\right| \le \left(|\varepsilon(h)| + T\left|\frac{\omega(h)}{h}\right|\right) e^{KT} \to 0, \qquad \text{as } h \to 0,
$$

which is the desired result. ∎

Let us now conclude the proof of Theorem 2.2.2. We know that

$$
\frac{\phi(x_h(T) - \phi(\bar{x}(T)}{h} \ge 0,
$$

ϕ is a C^1 function and for $t \in [s,T]$ we can write $x_h(t) = \bar{x}(t) + h\xi(t) + o(h,t)$, where $\sup_{[s,T]} \frac{|o(h,t)|}{h} \to 0$ as $h \to 0$. Then we obtain that

$$0 \leq \frac{\phi(\bar{x}(T) + h\xi(T) + o(h,T)) - \phi(\bar{x}(T))}{h} \to D\phi(\bar{x}(T)) \cdot \xi(T).$$

But if $M(t,s)$ is the fundamental solution (in s) of the equation of ξ, that is the unique solution of

$$\begin{cases} M'(t,s) = D_x f(t, \bar{x}(t), \bar{u}(t)) M(t,s) \quad \text{on } (s,T], \\ \\ M(s,s) = I, \end{cases}$$

then ξ is given by $\xi(t) = M(t,s)\,[f(s, x_h(s), v) - f(s, \bar{x}(s), \bar{u}(s))]$ and the previous inequality becomes

$$\langle D\phi(\bar{x}(T)), M(T,s) f(s, x_h(s), v)\rangle \geq \langle D\phi(\bar{x}(T)), M(T,s) f(s, \bar{x}(s), \bar{u}(s))\rangle,$$

or, equivalently,

$$\langle M^*(T,s) D\phi(\bar{x}(T)), f(s, x_h(s), v)\rangle \geq \langle M^*(T,s) D\phi(\bar{x}(T)), f(s, \bar{x}(s), \bar{u}(s))\rangle,$$

The proof is then complete, once we observe that the solution of equation (2.13) is given by $\lambda(s) = M^*(T,s) D\phi(\bar{x}(T))$, where $M^*(T,s)$ is the adjoint of $M(T,s)$. Indeed, the semigroup property of the fundamental matrix gives

$$M(t,r)M(r,s) = M(t,s) \qquad \text{for any } s \leq r \leq t.$$

Then,

$$\frac{\partial M}{\partial s}(t,r) M(r,s) + M(t,r) \frac{\partial M}{\partial t}(r,s) = 0$$

and for $r = s$ we get

$$\frac{\partial M}{\partial s}(t,s) = -M(t,s) D_x f(s, \bar{x}(s), \bar{u}(s)).$$

This means that

$$\frac{\partial M^*}{\partial s}(t,s) = -(D_x f)^*(s, \bar{x}(s), \bar{u}(s)) M^*(t,s),$$

which is the equation of the fundamental matrix related to equation (2.13). ■

Remark 2.2.5 If our control problem is a maximization one, then, in the above reasoning, everything stays the same except for the direction of inequalities; so we have the so called *Pontryagin Maximum Principle*

$$\langle \lambda(t), f(t, \bar{x}(t), \bar{u}(t))\rangle = \max_{v \in U} \langle \lambda(t), f(t, \bar{x}(t), v)\rangle.$$

In the case of a Bolza problem, the Pontryagin Minimum Principle can be formulated as follows:

Theorem 2.2.6 (Pontryagin Minimum Principle) *Assume that $\phi \in C^1(\mathbb{R}^n)$, $D_x f$ is continuous on $[0,T] \times \mathbb{R}^n \times U$ and that the Lagrangian L satisfies the same assumptions of f, that is (2.4) and the continuity of $D_x L$. Let $\{\bar{u}, \bar{x}\}$ be an optimal pair for problem (2.5)–(2.6) and let $\lambda(\cdot)$ be the solution of the adjoint equation*

$$\begin{cases} -\lambda'(t) = D_x L(t, \bar{x}(t,\bar{u}), \bar{u}(t)) + (D_x f)^* (t, \bar{x}(t,\bar{u}), \bar{u}(t)) \cdot \lambda(t) \\ \\ \lambda(T) = D\phi(\bar{x}(T,\bar{u})). \end{cases} \tag{2.16}$$

Then,

$$\langle \lambda(t), f(t, \bar{x}(t), \bar{u}(t)) \rangle + L(t, \bar{x}(t), \bar{u}(t)) = \min_{v \in U} \langle \lambda(t), f(t, \bar{x}(t), v) \rangle + L(t, \bar{x}(t), v) \tag{2.17}$$

in any continuity point t of $\bar{u}$.

Proof– We have seen in Remark 2.1.4 that we can reformulate a Bolza problem (2.5)–(2.6) into a Mayer one (2.8)–(2.7). Hence recalling the definitions of X, F, Φ in Remark 2.1.4, and denoting by λ_z, λ, the co–state variable of the state variable z, x respectively, Theorem 2.2.2 gives

$$\langle \Lambda(t), F(t, \bar{X}(t), \bar{u}(t)) \rangle = \min_{v \in U} \langle \Lambda(t), F(t, \bar{X}(t), v) \rangle,$$

where Λ is the solution of the co–state equation

$$\begin{cases} -\Lambda'(t) = (D_X F)^* \left(t, \bar{X}(t), \bar{u}(t)\right) \Lambda(t) \\ \\ \Lambda(T) = D\Phi(\bar{X}(T)). \end{cases} \tag{2.18}$$

Here

$$i) \quad D_X F = \left(\frac{\partial F_i}{\partial X_j} \right)_{i,j=1\ldots n} = \begin{pmatrix} 0 & D_x L \\ \\ 0 & D_x f \end{pmatrix}$$

$$ii) \quad D\Phi = \begin{pmatrix} 1 \\ \\ D\phi \end{pmatrix}$$

$$iii) \quad \Lambda(t) = \begin{pmatrix} \lambda_z(t) \\ \\ \lambda(t) \end{pmatrix}.$$

Writing (2.18) in vector form, we obtain

$$\begin{cases} -\lambda_z'(t) = 0 \\ \\ -\lambda'(t) = D_x L(t, \bar{x}(t,\bar{u}), \bar{u}(t)) \lambda_z(t) + (D_x f)^* (t, \bar{x}(t,\bar{u}), \bar{u}(t)) \cdot \lambda(t) \\ \\ \lambda_z(T) = 1 \\ \\ \lambda(T) = D\phi(\bar{x}(T)), \end{cases}$$

which yields (2.16). ∎

Let $\mathcal{H} = \mathcal{H}(t, x, v, \lambda)$ be the pre–Hamiltonian function defined as

$$\mathcal{H}(t, x, v, \lambda) =< \lambda, f(t, x, v) > +L(t, x, v)$$

and associated with the optimal control problem (2.5)–(2.6). Then necessary optimality conditions can be restated as follows:

Theorem 2.2.7 *Assume hypotheses of Theorem 2.2.6. Let $\{\bar{u}, \bar{x}\}$ be an optimal pair for problem (2.5)–(2.6) and let λ be the corresponding co–state. Then, λ solves*

$$\mathcal{H}(t, \bar{x}(t), \bar{u}(t), \lambda(t)) = \min_{v \in U} \mathcal{H}(t, \bar{x}(t), v, \lambda(t)), \tag{2.19}$$

in any continuity point t of $\bar{u}$.
Moreover, $\bar{x}$ and λ can be characterized as the solutions of the Hamiltonian system

$$\begin{cases} \lambda'(t) = -\dfrac{\partial \mathcal{H}}{\partial x}(t, \bar{x}(t), \bar{u}(t), \lambda(t)) & \lambda(T) = D\phi(\bar{x}(T)) \\ \\ \bar{x}'(t) = \dfrac{\partial \mathcal{H}}{\partial \lambda}(t, \bar{x}(t), \bar{u}(t), \lambda(t)) & \bar{x}(0) = x_0. \end{cases} \tag{2.20}$$

Remark 2.2.8 Let us consider a Bolza maximization problem,

$$\max_{u(\cdot)} \int_0^T L(t, x(t), u(t))\, dt + \phi(x(T))$$

where $x = x(t; x_0, u)$ is a solution of the controlled system (2.6). Remarking that a function f has a maximum at x if and only if $-f$ has a minimum at x and reviewing Theorem 2.2.6, we obtain that the co–state equation related to a Bolza maximization problem is still (2.16) or (2.20), while the Pontryagin principle turns out to be a maximization condition, i.e.:

$$\langle \lambda(t), f(t, \bar{x}(t), \bar{u}(t)) \rangle + L(t, \bar{x}(t), \bar{u}(t)) = \max_{v \in U} \langle \lambda(t), f(t, \bar{x}(t), v) \rangle + L(t, \bar{x}(t), v).$$

(Pontryagin Maximum Principle)

Recalling that $\mathcal{H}(t, x, u, \lambda) =< \lambda(t), f(t, x, u) > +L(t, x, v)$ we can rewrite the above principle as,

$$\mathcal{H}(t, \bar{x}(t), \bar{u}(t), \lambda(t)) = \max_{v \in U} \mathcal{H}(t, \bar{x}(t), v, \lambda(t)).$$

Very often, in applications of Optimal Control Theory to economics, the Lagrange function is taken of the form

$$L(t, x, v) = F(t, x, v) e^{-\rho t},$$

where ρ is a positive constant. For this kind of problems, instead of considering the co–state λ and the pre–Hamiltonian function $\mathcal{H}$, as we did so far, it is sometimes useful to introduce the so called *current value co–state* λ_c and the *current value pre–Hamiltonian* $\mathcal{H}_c$ as follows, see [5, p. 210]:

$$\lambda_c = \lambda e^{\rho t}, \qquad \mathcal{H}_c(t, x, v, \lambda_c) = \mathcal{H}(t, x, v, \lambda) e^{\rho t}.$$

Let us consider the maximization problem,

$$\max_{u(\cdot)} \int_0^T F(t, x(t), u(t)) e^{-\rho t} dt + \phi_T(x(T)),$$

where $x = x(t; x_0, u)$ is a solution of the controlled system (2.6). If $\{\bar{x}(\cdot), \bar{u}(\cdot)\}$ is an optimal pair for such a problem, then the above theorem can be reformulated in terms of λ_c and $\mathcal{H}_c$ as follows:

$$\mathcal{H}_c(t, \bar{x}(t), \bar{u}(t), \lambda_c(t)) = \max_{v \in U} \mathcal{H}_c(t, \bar{x}(t), v, \lambda_c(t))$$

and $\bar{x}$ and λ_c solve the Hamiltonian system

$$\begin{cases} \lambda_c'(t) = -\dfrac{\partial \mathcal{H}_c}{\partial x}(t, \bar{x}(t), \bar{u}(t), \lambda_c(t)) + \rho \lambda_c(t) & \lambda_c(T) = D\phi(\bar{x}(T)) e^{\rho t} \\ \\ \bar{x}'(t) = \dfrac{\partial \mathcal{H}_c}{\partial \lambda_c}(t, \bar{x}(t), \bar{u}(t), \lambda_c(t)) & \bar{x}(0) = x_0. \end{cases}$$

2.2.1 Examples

In this section, we will present three applications of the Pontryagin Minimum Principle to economics and we will also discuss the introductory Examples 2.1.3 and 2.1.5. For the reader's convenience, we begin with a simple mathematical example of the Mayer optimal control problem in order to see how the principle applies.

Example 2.2.9 Let $x_0 \in \mathbb{R} \setminus \{0\}$, $T = 1$ and $U = [-1, 1]$. The problem is

$$\max_{u(\cdot)} \frac{1}{2} x^2(1, u) \tag{2.21}$$

where $u : [0, 1] \to [-1, 1]$ and $x = x(t, u)$ is a solution of the controlled system

$$\begin{cases} x'(t) = x(t) u(t) \\ \\ x(0) = x_0 \neq 0. \end{cases} \tag{2.22}$$

Optimal trajectories are computed by hand

$$\bar{x}(t) = x_0 e^t. \tag{2.23}$$

If we solve the problem via the Pontryagin Maximum Principle we get

$$\lambda(t) \bar{u}(t) \bar{x}(t) = \max_{-1 \le u \le 1} \lambda(t) u \, \bar{x}(t) = |\lambda(t) \bar{x}(t)| \quad \text{and} \quad \bar{u}(t) = \text{sgn}(\lambda(t) \bar{x}(t)).$$

On the other hand

$$\begin{cases} \lambda'(t) = \bar{u}(t)\lambda(t) \\ \\ \lambda(1) = \bar{x}(1); \end{cases}$$

hence

$$\lambda(t) = \bar{x}(1)e^{\int_t^1 \bar{u}(s)\ ds}.$$

Since both $\lambda(t)$ and $\bar{x}(t)$ have the same sign as $\bar{x}(1)$, we conclude that $\bar{u}(t) = 1$. Setting the optimal control in the state equation we obtain the optimal trajectories $\bar{x}(t) = x_0 e^t$.

Economic applications of Optimal Control Theory raised between 1965 and 1975. In the sequel we shall present examples both in macro and in microeconomic theory. In the following example we present a model which shows how to choose an optimal portfolio in the presence of transaction cost, see [14, p. 254].

Example 2.2.10 An agent has a capital to invest in a remunerative bank account or in a stock. The setup is deterministic and we assume to know the instantaneous remuneration rate $r(t)$ of the bank account, the dividend flow $\delta(t)$ associated with the stock and the stock prize $S(t)$, for each $t \in [0, T]$. At any time t we can buy or sell a maximum number of stocks M and for every transaction we pay a cost $\alpha \in (0, 1)$. We denote by $x_1(t)$ the amount of money on the bank account at time t and by $x_2(t)$ the number of stocks in the portfolio at time t. Moreover we call $v(t)$ the algebraic number of stocks sold at time t. We are interested in maximizing the portfolio value at T that is

$$\max_{v(\cdot)}\{x_1(T) + S(T)x_2(T)\} \tag{2.24}$$

subject to

$$\begin{cases} x_1'(t) = r(t)x_1(t) + \delta(t)x_2(t) + S(t)(v(t) - \alpha|v(t)|) \\ \\ x_2'(t) = -v(t) \\ \\ x_1(0) = x_1 \\ \\ x_2(0) = x_2 \end{cases} \tag{2.25}$$

and $v : [0, T] \to [-M, M]$. Let $\bar{u}$ be an optimal control and $(\bar{x}_1, \bar{x}_2)$ the related optimal trajectory. The co–state equations are

$$\begin{cases} -\lambda_1'(t) & = & r(t)\lambda_1(t) \\ \\ -\lambda_2'(t) & = & \delta(t)\lambda_1(t) \\ \\ \lambda_1(T) & = & 1 \\ \\ \lambda_2(T) & = & S(T), \end{cases} \tag{2.26}$$

whose solutions are

$$\lambda_1(t) = e^{\int_t^T r(\tau)d\tau} \quad \text{and} \quad \lambda_2(t) = S(T) + \int_t^T \delta(\tau) e^{\int_\tau^T r(s)ds} d\tau.$$

Applying the Pontryagin Maximum Principle, we also have

$$\begin{aligned} &\lambda_1(t)[r(t)\bar{x}_1(t) + \delta(t)\bar{x}_2(t) + S(t)(\bar{u}(t) - \alpha|\bar{u}(t)|)] - \lambda_2(t)\bar{u}(t) \\ &= \max_{|v|\le M} \{\lambda_1(t)[r(t)x_1(t) + \delta(t)x_2(t) + S(t)(v - \alpha|v|)] - \lambda_2 v\}, \end{aligned}$$

which gives

$$\lambda_1(t)S(t)(\bar{u}(t) - \alpha|\bar{u}(t)|) - \lambda_2(t)\bar{u}(t)] = \max_{|v|\le M} \{v\,(\lambda_1(t)S(t)(1 - \alpha\mathrm{sgn}(v)) - \lambda_2(t)v)\}.$$

Exploiting the previous equation we find that the optimal control $\bar{u}$ is given by

$$\bar{u}(t) = \begin{cases} M & \text{if} \quad \lambda_1(t)S(t)(1-\alpha) > \lambda_2(t); \\ 0 & \text{if} \quad \lambda_1(t)S(t)(1-\alpha) < \lambda_2(t) < \lambda_1(t)S(t)(1+\alpha); \\ -M & \text{if} \quad \lambda_1(t)S(t)(1+\alpha) < \lambda_2(t). \end{cases} \tag{2.27}$$

Denoting by $q(t)$ the ratio $\dfrac{\lambda_2(t)}{S(t)\lambda_1(t)}$ the control $\bar{u}$ can also be written as

$$\bar{u}(t) = \begin{cases} M & \text{if} \quad q(t) < 1-\alpha; \\ 0 & \text{if} \quad 1-\alpha < q(t) < 1+\alpha; \\ -M & \text{if} \quad q(t) > 1+\alpha. \end{cases}$$

Once we specify the data $r(\cdot)$, $\delta(\cdot)$, $S(\cdot)$ and α we can say for which t $\bar{u}$ is equal to $-M$, 0 or M.

The next example is due to Nordhaus in 1975. He discussed the political business cycles which arise in democratic countries within each electoral period, see [18] or [5, p. 193].

Example 2.2.11 In a democratic nation there are elections to decide which party is going to rule the country. In the Nordhaus Model, the party which is in charge of the government wishes to get more voters considering only two economic variables: u, the unemployment rate, and π, the inflation rate. These variables are considered in a vote function $V = V(u, \pi)$ which represents how many votes the ruling party is going to get managing a policy which acts on the unemployment rate and on the inflation rate. V is assumed to be a decreasing function with respect to the u and π variables:

$$\frac{\partial V}{\partial u} < 0 \text{ and } \frac{\partial V}{\partial \pi} < 0,$$

that is there is an interplay between inflation and unemployment.
Moreover u and π are related through the Phillips tradeoff relation

$$p(t) = \phi(u(t)) + a\pi(t),$$

where ϕ is a decreasing function, a is a constant such that $a \in (0, 1]$ and π is the expected rate of inflation which behaves accordingly to the adaptive differential equation

$$\pi'(t) = b(p(t) - \pi(t)), \qquad \text{b positive constant.}$$

Assuming T to be the time left to the elections'day and taking into account the Phillips relation, the problem that the incumbent party has to face is the following: to maximize on the controls u the functional

$$\int_0^T V(u(t), \phi(u(t)))e^{rt}dt,$$

where $r > 0$ and e^{rt} is a decay memory factor (voters give more importance to what the government does near the elections' day).
Again, according to the Phillips relation the state equation is

$$\begin{cases} \pi'(t) = b\phi(u(t)) + b(a-1)\pi(t) \\ \\ \pi(0) = \pi_0. \end{cases}$$

The aim is to find the optimal inflation rate $\bar{u}$ the government has to implement. Although u cannot be negative, there are no constraints on the sign of u. In the Nordhaus framework

$$\phi(u) = j - ku \qquad \text{and} \qquad V(u, \pi) = -(u^2 + hp), \quad j, k, h > 0.$$

The optimal control problem is to maximize

$$-\int_0^T [u^2(t) + h(j - ku(t) + a\pi(t))]e^{rt}\ dt,$$

which is equivalent to minimize

$$\int_0^T [u^2(t) + h(j - ku(t) + a\pi(t))]e^{rt}\ dt,$$

subject to

$$\begin{cases} \pi'(t) = b[j - ku(t) + (a-1)\pi(t)] \\ \\ \pi(0) = \pi_0. \end{cases}$$

The pre–Hamiltonian function of this problem is

$$\mathcal{H}(t, \pi, u, \lambda) = e^{rt}[u^2 + h(j - ku + a\pi)] + \lambda b[j - ku + (a-1)\pi].$$

If $\bar{u}$ is an optimal control, then it minimizes $\mathcal{H}$ and hence we have

$$0 = \frac{\partial \mathcal{H}}{\partial u} = 2ue^{rt} - hke^{rt} - bk\lambda.$$

Since

$$\frac{\partial^2 \mathcal{H}}{\partial u^2} = 2e^{rt} > 0,$$

$\mathcal{H}$ attains its unique minimum at

$$\bar{u}(t) = \frac{hk}{2} + \frac{bk}{2}\bar{\lambda}(t)e^{-rt}.$$

So, knowing $\bar{\lambda}$ we have the optimal control.
In order to find $\bar{\lambda}$, we solve the co–state equation

$$\begin{cases} \lambda'(t) = -\dfrac{\partial \mathcal{H}}{\partial \pi}(t, \bar{\pi}(t), \bar{u}(t), \lambda(t)) = ahe^{rt} + b(a-1)\lambda(t) \\ \\ \lambda(T) = 0. \end{cases}$$

The solution of the above Cauchy problem can be easily obtained:

$$\bar{\lambda}(t) = \frac{ah}{r + b(a-1)}[e^{[r+b(a-1)]T+(1-a)bt} - e^{rt}].$$

Finally, substituting $\bar{\lambda}$ into $\bar{u}$ we get the optimal unemployment rate

$$\begin{aligned} \bar{u}(t) &= \frac{hk}{2} + \frac{bk}{2}\frac{ah}{r + b(a-1)}[e^{[r+b(a-1)](T-t)} - 1] \\ &= \frac{hk}{2[r + b(a-1)]}\left[r - b + abe^{[r+b(a-1)](T-t)}\right]. \end{aligned}$$

We remark that since

$$\bar{u}'(t) = -\frac{hk}{2}abe^{[r+b(a-1)](T-t)} < 0,$$

$\bar{u}$ is a decreasing function of t. Moreover, since $\bar{u}(T) = \frac{hk}{2} > 0$, we conclude that the optimal unemployment rate is always positive.

Our last economic example is the so–called Eisner–Strotz model to expand a firm's plant size. We will analyse such a model by means of the current value pre–Hamiltonian formulation.

Example 2.2.12 The Eisner–Strotz model analyses the problem of a firm that wants to expand the machinary used in its industrial process in a fixed interval of time $[0, T]$, $T > 0$. The profit rate π associated with each plant size is supposed to be a known function and it will be evaluated on the capital stock K, which is a measure of the plant size. In the expansion process the firm has to face an adjustment cost C, which depends on the velocity K' of the expansion. On the other hand, the capital stock K satisfies the equation $K'(t) = I(t)$, where $I(t)$ is the net investment and our control variable. Indeed, the more the firm invests, the fastest the expansion is. The optimal control problem we are dealing with is

$$\max_{I(\cdot)} \int_0^T [\pi(K(t)) - C(I(t))]\, e^{-\rho t}\, dt,$$

where $\rho > 0$, $I : [0,T] \to \mathbb{R}$ and K is the solution of the controlled system

$$\begin{cases} K'(t) = I(t) \\ \\ K(0) = K_0, \qquad K_0 > 0. \end{cases}$$

Roughly speaking, the above maximization problem is the mathematical representation of the firm's aim of finding the optimal path $K^*(t)$ which maximizes the total discounted value of its net profit (see [10] or [5, p. 106] for further details). In order to give quantitative solutions, we assume that

$$\pi(K) = \alpha K - \beta K^2 \qquad \text{and} \qquad C(I) = aI^2 + bI,$$

where α, β, a and b are given positive constants. In this case, the current value pre–Hamiltonian is

$$\mathcal{H}_c(t, K, I, \lambda_c) = \alpha K - \beta K^2 - aI^2 - bI + \lambda_c I,$$

whose maximum value with respect to the control variable I is attained at

$$I = \frac{\lambda_c - b}{2a}.$$

Hence, the Hamiltonian system is

$$\begin{cases} \lambda_c'(t) = -\alpha + 2\beta K(t) + \rho\lambda_c(t), \quad \lambda_c(T) = 0 \\ \\ K'(t) = \dfrac{\lambda_c(t) - b}{2a}, \quad K(0) = K_0. \end{cases}$$

Solving the above system, we can easily derive the required optimal path

$$K^*(t) = c_1 e^{r_1 t} + c_2 e^{r_2 t} + \bar{K},$$

where

$$\begin{cases} \bar{K} = \dfrac{\alpha - \rho b}{2\beta} \\ \\ r_1 = \dfrac{\rho a + \sqrt{\rho^2 a^2 + 4a\beta}}{2a} \qquad r_2 = \dfrac{\rho a - \sqrt{\rho^2 a^2 + 4a\beta}}{2a} \\ \\ c_1 = \dfrac{(K_0 - \bar{K}) r_2 e^{r_2 T} + b/2a}{r_2 e^{r_2 T} - r_1 e^{r_1 T}} \qquad c_2 = -\dfrac{(K_0 - \bar{K}) r_1 e^{r_1 T} + b/2a}{r_2 e^{r_2 T} - r_1 e^{r_1 T}}. \end{cases}$$

Let us now discuss Examples 2.1.3 and 2.1.5, starting from the model of the boat entering in a basin.

Example 2.2.13 Let $T > 0$ be fixed and $U = \{(u_1, u_2) \mid u_1^2 + u_2^2 \le 1\}$. The problem is

$$\min \ -x_1(T),$$

where $(x_1, x_2) : [0, T] \to \mathbb{R}^2$ is a solution of the system

$$\begin{cases} x_1'(t) = x_2(t) + u_1(t) \qquad t \in (0, T) \\ x_2'(t) = u_2(t) \qquad t \in (0, T) \\ x_1(0) = 0 \\ x_2(0) = 0, \end{cases}$$

and $(u_1, u_2) \in U$.
Consider the optimal pair $\{\bar{x}(t), \bar{u}(t)\}$. Then the co–state equation is

$$\begin{cases} \lambda_1'(t) = 0, \\ \lambda_2'(t) = -\lambda_1(t) \\ \lambda_1(T) = -1, \\ \lambda_2(T) = 0, \end{cases}$$

which gives $\lambda_1(t) = -1$ and $\lambda_2(t) = t - T$. Applying the Pontryagin Minimum Principle we have that

$$\min_{u \in U}\{\lambda_1(t)(\bar{x}_2(t) + u_1) + \lambda_2(t)u_2\} = \lambda_1(t)\bar{x}_2(t) + \min_{u \in U}\{\lambda_1(t)u_1 + \lambda_2(t)u_2\}$$

is attained at

$$\bar{u}_1(t) = -\frac{\lambda_1(t)}{\sqrt{\lambda_1^2(t) + \lambda_2^2(t)}} \qquad \text{and} \qquad \bar{u}_2(t) = -\frac{\lambda_2(t)}{\sqrt{\lambda_1^2(t) + \lambda_2^2(t)}},$$

i.e.

$$\bar{u}_1(t) = \frac{1}{\sqrt{1 + (t - T)^2}} \qquad \text{and} \qquad \bar{u}_2(t) = \frac{T - t}{\sqrt{1 + (t - T)^2}}.$$

By inserting the optimal control $\bar{u}$ into the state equation we obtain

$$\begin{cases} x_1'(t) = x_2(t) + \dfrac{1}{\sqrt{1+(t-T)^2}} & t \in (0,T) \\[2ex] x_2'(t) = \dfrac{T-t}{\sqrt{1+(t-T)^2}} & t \in (0,T) \\[2ex] x_1(0) = 0 \\[2ex] x_2(0) = 0, \end{cases}$$

which produces the optimal trajectory

$$\begin{cases} \bar{x}_1(t) = -\dfrac{1}{2}(t-T)\sqrt{1+(t-T)^2} - \dfrac{T}{2}\sqrt{1+T^2} + t\sqrt{1+T^2} \\[2ex] \qquad\quad -\dfrac{1}{2}\log\left(T - t + \sqrt{1+(t-T)^2}\right) + \dfrac{1}{2}\log\left(T + \sqrt{1+T^2}\right) \\[2ex] \bar{x}_2(t) = -\sqrt{1+(t-T)^2} + \sqrt{1+T^2} \end{cases}$$

Let us now complete the discussion of the model of pumping out water from a basin, see Example 2.1.5.

Example 2.2.14 Let $T > 0$ be fixed and $U = [0,1]$. We have already seen that the problem is

$$\min_{u(\cdot)} \int_0^T u^2(t)dt + x^2(T), \tag{2.28}$$

where $u : [0,T] \to [0,1]$ and $x = x(t;u)$ is a solution of the controlled system

$$\begin{cases} x'(t) = -u(t) \\[2ex] x(0) = 1. \end{cases} \tag{2.29}$$

Consider the optimal pair $\{\bar{x}(t), \bar{u}(t)\}$. Then the co–state equation is

$$\begin{cases} \lambda'(t) = 0, \\[2ex] \lambda(T) = 2\bar{x}(T), \end{cases}$$

which gives $\lambda(t) = 2\bar{x}(T)$. Applying the Pontryagin Minimum Principle we have that

$$\min_{v \in U} \langle \lambda(t), -v \rangle + v^2$$

is attained at $\bar{u}(t) \equiv \bar{x}(T)$. By inserting the optimal control $\bar{u}$ into the state equation

$$\begin{cases} \bar{x}'(t) = -\bar{x}(T) \\ \\ \bar{x}(0) = 1, \end{cases}$$

we obtain $\bar{x}(t) = 1 - t\bar{x}(T)$ and therefore $\bar{x}(T) = \dfrac{1}{1+T}$. Hence the optimal control and the optimal trajectory are respectively

$$\bar{u}(t) = \frac{1}{1+T} \quad \text{and} \quad \bar{x}(t) = 1 - \frac{t}{1+T}.$$

2.2.2 Time Optimal Problem

In the optimal control theory we introduced so far the interval $[0, T]$ was fixed. Then we cannot directly apply it to problems where the final time varies with the control variable. However, using the previous computations, we can formulate a suitable Minimum Principle for the Time Optimal Problem. Let us first introduce the general setting for Optimal Control Problems with Exit Time.

Let $C \in \mathbb{R}^n$ be a closed nonempty set, and take $U \in \mathbb{R}^k$, $f \in C([0,T] \times \mathbb{R}^n \times U)$, $L : [0,T] \times \mathbb{R}^n \times U \to \mathbb{R}$ and $\phi : \mathbb{R}^n \to \mathbb{R}$ as in Definition 2.1.2. Moreover, for any $x_0 \in \mathbb{R}^n \setminus C$ and $u \in C_*([0,\infty), U)$ define the (extended valued) map

$$\tau(x_0, u) = \inf\{t > 0 : x^{x_0}(t, u) \in C\},$$

where $x^{x_0}(\cdot, u) \in C^1_*([0,T], \mathbb{R}^n)$ is the solution of

$$\begin{cases} x'(t) = f(t, x(t), u(t)) \\ \\ x(0) = x_0. \end{cases} \tag{2.30}$$

Then the Optimal Control Problem with Exit Time is the following

$$\inf_{u(\cdot) \in C_*([0,\infty),U)} \left\{ \int_0^{\tau(x_0,u)} L(t, x^{x_0}(t,u), u(t))\, dt + \phi(x^{x_0}(\tau(x_0,u), u)) \right\}, \tag{2.31}$$

with $x^{x_0}(\cdot, u)$ and $\tau(x_0, u)$ as above.
In the language of Optimal Control Theory, the set C is called *target* and the map τ is called *transition time*.

In this section we are interested in the particular problem with $\phi \equiv 0$ and $L \equiv 1$, which is the *Time Optimal Problem*:

$$\inf_{u(\cdot) \in C_*([0,\infty),U)} \tau(x_0, u), \tag{2.32}$$

where $\tau(x_0, u) = \inf\{t > 0 : x^{x_0}(t, u) \in C\}$ and $x^{x_0}(\cdot, u) \in C^1_*([0, T], \mathbb{R}^n)$ is the solution of

$$\begin{cases} x'(t) = f(t, x(t), u(t)) \\ \\ x(0) = x_0. \end{cases} \tag{2.33}$$

Theorem 2.2.15 *Assume that there exists $d \in C^1(\mathbb{R}^n)$ such that $C = \{x \in \mathbb{R}^n \ : \ d(x) \leq 0\}$ and that $\nabla d \neq 0$ on the set $\partial C = \{x \in \mathbb{R}^n \ : \ d(x) = 0\}$. Let $\{\bar{u}, \bar{x}\}$ be an optimal pair for problem (2.32) (here $\bar{x}(\cdot)$ stands for $x^{x_0}(\cdot, \bar{u})$), i.e.*

$$\tau(x_0, \bar{u}) = \min_{u(\cdot) \in C_*([0,\infty), U)} \tau(x_0, u)$$

and $\bar{x}$ solves (2.33). Then the solution of the adjoint equation (called co–state*)*

$$\begin{cases} -\lambda'(t) = (D_x f)^* (t, \bar{x}(t), \bar{u}(t)) \cdot \lambda(t) \\ \\ \lambda(T) = \nu(\bar{x}(\tau(x_0, \bar{u}))), \end{cases} \tag{2.34}$$

where $\nu(\bar{x}(\tau(x_0, \bar{u})))$ is the outer unit normal to C in $\bar{x}(\tau(x_0, \bar{u}))$, satisfies the Pontryagin Minimum Principle

$$\langle \lambda(t), f(t, \bar{x}(t), \bar{u}(t)) \rangle = \min_{v \in U} \langle \lambda(t), f(t, \bar{x}(t), v) \rangle \tag{2.35}$$

in any continuity point $t \in (0, \tau(x_0, \bar{u})]$ of the optimal control $\bar{u}$.

Proof– Let us call $\bar{\tau}$ the optimal arrival time $\tau(x_0, \bar{u})$. Fix any continuity point $s \in (0, \bar{\tau}]$ of $\bar{u}$, $h \in (0, s)$, $v \in U$ and define

$$u_h(t) = \begin{cases} v & \text{if } s - h < t \leq s \\ \\ \bar{u}(t) & \text{elsewhere }. \end{cases}$$

We denote by $x_h(\cdot)$ the corresponding state $x_h(\cdot, u_h)$. Arguing as in Theorem 2.2.2 we find that

$$\frac{x_h(t) - \bar{x}(t)}{h} \to \xi(t), \qquad \text{uniformly on } [s, \bar{\tau}],$$

where

$$\xi(t) = \begin{cases} 0 & \text{on } [0, s] \\ \\ M(t, s)[f(s, \bar{x}(s), v) - f(s, \bar{x}(s), \bar{u}(s))] & \text{on } (s, \bar{\tau}] \end{cases}$$

and $M(t, s)$ solves

$$\begin{cases} \dfrac{\partial M}{\partial t}(t, s) = D_x f(t, \bar{x}(t), \bar{u}(t)) M(t, s) \quad \text{on } (s, T], \\ \\ M(s, s) = I \end{cases}$$

Moreover, by the minimality of $\bar{\tau}$, we have

$$d(\bar{x}(\bar{\tau})) = 0 \qquad \text{and} \qquad d(x_h(\bar{\tau})) \geq 0.$$

Hence we find again that

$$0 \leq \frac{d(x_h(\bar{\tau})) - d(\bar{x}(\bar{\tau}))}{h} \to \langle \nabla d(\bar{x}(\bar{\tau})), \xi(\bar{\tau}) \rangle, \qquad \text{as } h \to 0,$$

i.e.

$$\langle \nabla d(\bar{x}(\bar{\tau})), M(\bar{\tau}, s)[f(s, \bar{x}(s), v) - f(s, \bar{x}(s), \bar{u}(s))] \rangle \geq 0.$$

Dividing by $|\nabla d(\bar{x}(\bar{\tau}))| \neq 0$ and taking into account that $\frac{\nabla d(\bar{x}(\bar{\tau}))}{|\nabla d(\bar{x}(\bar{\tau}))|}$ is precisely the outer unit normal to C in $\bar{x}(\bar{\tau})$ we obtain the assertion. ∎

Theorem 2.2.16 *Assume that $C = \{0\}$ and let $\{\bar{u}, \bar{x}\}$ be an optimal pair for problem (2.32) (here $\bar{x}(\cdot)$ stands for $x^{x_0}(\cdot, \bar{u})$). Then there exists some $\mu \in \mathbb{R}^n$ with $|\mu| = 1$ such that the solution of the adjoint equation*

$$\begin{cases} -\lambda'(t) = (D_x f)^* (t, \bar{x}(t), \bar{u}(t)) \cdot \lambda(t) \\ \\ \lambda(T) = \nu, \end{cases} \tag{2.36}$$

satisfies the Pontryagin Minimum Principle

$$\langle \lambda(t), f(t, \bar{x}(t), \bar{u}(t)) \rangle = \min_{v \in U} \langle \lambda(t), f(t, \bar{x}(t), v) \rangle \tag{2.37}$$

in any continuity point $t \in (0, \tau(x_0, \bar{u})]$ of the optimal control $\bar{u}$.

Proof– We will only give an idea of the proof. For any $r > 0$ take the perturbed optimal control problem with target $C_r = \{x \in \mathbb{R}^n \ : \ |x|^2 - r^2 \leq 0\}$. Let $\{u_r, x_r\}$ be an optimal pair, τ_r the optimal arrival time to C_r and λ_r the co–state. Then

$$\lambda_r(\tau_r) = \frac{x_r(\tau_r)}{|x_r(\tau_r)|}$$

is a unit vector for any $r > 0$. By taking the limit as $r \to 0$, it can be shown that $\{u_r, x_r\}$, τ_r, C_r and λ_r converge (by subsequence and in some sense to be specified) to $\{\bar{u}, \bar{x}\}$, $\bar{\tau}$, C and λ, respectively. Hence the vector $\lambda_r(\tau_r)$ converges to some unit vector which is the final datum for the co–state of problem (2.32) with target $C = \{0\}$. ∎

Remark 2.2.17 If the state equation (2.33) is autonomous, then it can be shown that actually in Theorems 2.2.15 and 2.2.16

$$\langle \lambda(t), f(\bar{x}(t), \bar{u}(t)) \rangle = 0,$$

along any optimal pair $\{\bar{u}, \bar{x}\}$ of the Time Optimal Problem (see [15], p. 143).

We present an example of a time–optimal problem whose state equation is a second order differential equation, see [5, p. 226].

Example 2.2.18 We study the time–optimal capture of a wandering particle. Let $x_0, y_0 \neq 0$ be fixed. The problem we face is to solve

$$\inf\{t \geq 0 \ : \ \text{there exists } u(\cdot) \in [-1,1] = U \text{ such that } x(t,u) = 0, x'(t,u) = 0\}$$

where $x(t,u)$ is subject to

$$\begin{cases} x''(t) = u(t), \\ x(0) = x_0, \\ x'(0) = y_0. \end{cases}$$

The control $u(\cdot)$ represents the force we apply to the particle to drive it to rest. Since we are considering a second order equation, introducing the new state variable $y(t) = x'(t)$ we can define

$$X(t) = \begin{pmatrix} x(t) \\ y(t) \end{pmatrix},$$

and the state equation above can be rewritten as a first order system

$$\begin{cases} X'(t) = \begin{pmatrix} x'(t) \\ y'(t) \end{pmatrix} = \begin{pmatrix} y(t) \\ u(t) \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} X(t) + \begin{pmatrix} 0 \\ 1 \end{pmatrix} u(t), \\ X(0) = \begin{pmatrix} x_0 \\ y_0 \end{pmatrix}. \end{cases}$$

Using this formulation, the initial problem reads as a time optimal problem in $\mathbb{R}^2$ with target $C = \{(0,0)\}$. The pre–Hamiltonian function $\mathcal{H}$ is

$$\begin{aligned} \mathcal{H}(x,y,u,\lambda,\mu) &= 1 + \begin{pmatrix} \lambda \\ \mu \end{pmatrix} \cdot \left[\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \end{pmatrix} u \right] \\ &= 1 + \begin{pmatrix} \lambda \\ \mu \end{pmatrix} \cdot \begin{pmatrix} y \\ u \end{pmatrix} = 1 + \lambda y + \mu u. \end{aligned}$$

The adjoint equations are

$$\begin{cases} \lambda'(t) = -\dfrac{\partial \mathcal{H}}{\partial x} = 0, \\ \mu'(t) = -\dfrac{\partial \mathcal{H}}{\partial y} = -\lambda(t). \end{cases}$$

Hence

$$\lambda(t) = c_1,$$

$$\mu(t) = c_2 - c_1 t,$$

where c_1 and c_2 are unknown constants such that $c_1^2 + (c_2 - c_1\tau(x_0,u))^2 = 1$.
The Pontryagin Minimum Principle yields

$$\bar{u}(t) = \text{argmin}\{1 + c_1\bar{y}(t) + (c_2 - c_1 t)u : u \in [-1,1]\}.$$

Since the pre–Hamiltonian function $\mathcal{H}$ is linear with respect to u we have

$$\bar{u}(t) = -\text{sgn}(c_2 - c_1 t) \in \{-1, 1\}$$

and then $\bar{u}$ is a bang–bang control, i.e. it takes value from one side of the boundary of U to the other.
Since c_1 and c_2 are not known there are four possible cases:
Case 1: Suppose $\mu > 0$. Then $\bar{u}(t) = -1$ and the optimal trajectories are

$$\begin{cases} \bar{x}(t) = x_0 + y_0 t - \frac{1}{2}t^2 \\ \bar{y}(t) = y_0 - t. \end{cases} \tag{2.38}$$

In the phase plane they are parabolas over which we move clockwise, as it can be seen by rewriting the optimal trajectories as

$$\bar{x}(t) = -\frac{1}{2}(y_0 - t)^2 + K_1 = -\frac{\bar{y}(t)^2}{2} + K_1, \text{ where } K_1 = x_0 + \frac{y_0^2}{2}$$

Case 2: Suppose $\mu < 0$. Then $\bar{u}(t) = 1$ and the optimal trajectories are

$$\begin{cases} \bar{x}(t) = x_0 + y_0 t + \frac{1}{2}t^2 \\ \bar{y}(t) = y_0 + t. \end{cases}$$

Again, in the phase plane, they are parabolas over which we move clockwise, as it can be seen by rewriting the optimal trajectories as

$$\bar{x}(t) = \frac{\bar{y}(t)^2}{2} + K_2, \text{ where } K_2 = x_0 - \frac{y_0^2}{2}$$

In the cases we analyse so far, in order to have $\bar{x}(T) = 0$ we need to start at some point on the parabolas passing through the origin which are the ones where $K_1 = 0$ in case 1 or $K_2 = 0$ in case 2. Since we are dealing with bang–bang solutions the optimal control $\bar{u}$ may switch from values on the boundary of the control space U.
We denote by Π^-, Π^+, Ω^- and Ω^+ the sets

$$\Pi^- = \left\{(x,y)|x = -\frac{y^2}{2},\ y \geq 0\right\}, \qquad \Pi^+ = \left\{(x,y)|x = \frac{y^2}{2},\ y \leq 0\right\},$$

$$\Omega^- = \left\{(x,y)|x < -\frac{y^2}{2},\ y>0\right\} \cup \left\{(x,y)|x < \frac{y^2}{2},\ y<0\right\}$$

and

$$\Omega^+ = \left\{(x,y)|x > -\frac{y^2}{2},\ y>0\right\} \cup \left\{(x,y)|x > \frac{y^2}{2},\ y<0\right\}.$$

We have two other cases.
Case 3: Suppose $(x_0, y_0) \in \Omega^+$. In this case we catch a parabola associated with $\bar{u}(t) = -1$, we keep using this optimal control for a time $\bar{t}$ at which we hit Π^+ at a certain point $(x_1, y_1) = (\bar{x}(\bar{t}), \bar{y}(\bar{t}))$. Hence (x_1, y_1) is such that

$$\begin{cases} x_1 = -\dfrac{y_1^2}{2} + \left(x_0 + \dfrac{y_0^2}{2}\right) \\ \\ x_1 = \dfrac{y_1^2}{2}. \end{cases}$$

Solving the system we obtain

$$(x_1, y_1) = \left(\frac{1}{2}\left(x_0 + \frac{y_0^2}{2}\right), -\sqrt{x_0 + \frac{y_0^2}{2}}\right).$$

Moreover from (2.38) we can evaluate $\bar{t}$:

$$y_1 = \bar{y}(\bar{t}) = y_0 - \bar{t};$$

therefore

$$\bar{t} = y_0 - y_1 = y_0 + \sqrt{x_0 + \frac{y_0^2}{2}},$$

which means that the time elapsed is equal to the distance swept by $\bar{y}$.
From time $\bar{t}$ on, we drive along on Π^+ using the control $\bar{u}(t) = 1$ until we hit the origin. Recalling that on Π^+ we have $0 = \bar{y}(\bar{t}) + t$, the time elapsed is

$$t = -\bar{y}(\bar{t}) = -y_1 = \sqrt{x_0 + \frac{y_0^2}{2}},$$

and the total driving time from (x_0, y_0) to the origin is

$$\tau(x_0, y_0) = \bar{t} + t = y_0 + 2 + \sqrt{x_0 + \frac{y_0^2}{2}}.$$

Case 4: Suppose $(x_0, y_0) \in \Omega^-$. In this case we catch a parabola associated with $\bar{u}(t) = 1$, we keep using this optimal control for a time $\bar{t}$ at which we hit Π^- at a certain point $(x_2, y_2) = (\bar{x}(\bar{t}), \bar{y}(\bar{t}))$. From time $\bar{t}$ on, we drive along on Π^- using the control $u^*(t) = -1$ until we hit the origin. The reasoning is the same as the one used in the previous case. It is easy to see that now

$$\tau(x_0, y_0) = -y_0 + 2 + \sqrt{\frac{y_0^2}{2} - x_0}.$$

Exercise 2.2.19 *Describe the level sets of the function τ in the previous example.*

Exercise 2.2.20 *Analyse the time–optimal capture of a wandering particle (see Example 2.2.18) when the target (in the phase plane) is the unit ball in $\mathbb{R}^2$.*

Exercise 2.2.21 *Analyse the time–optimal capture of a wandering particle (see Example 2.2.18) when the target (in the phase plane) is the unit square in $\mathbb{R}^2$.*

Exercise 2.2.22 *Calculate the transition time $\tau(x_0, y_0)$, $(x_0, y_0) \in \mathbb{R}^2$, related to the equation*

$$\begin{cases} x'(t) = u(t) \\ \\ y'(t) = v(t) \\ \\ x(0) = x_0 \\ \\ y(0) = y_0 \end{cases}$$

where $|u|, |v| \leq 1$ and the target is $C = \{(0,0)\}$.

The next example concerns the smooth landing (with velocity zero) of a space vehicle on the surface of a planet along a vertical trajectory, see [15].

Example 2.2.23 (Soft Landing Problem) Denote by $x(t)$ the height at time t, $y(t)$ the instantaneous velocity and $z(t)$ the total mass of the vehicle (which is a nonincreasing function of time, since fuel is being consumed). If we denote by $u(t)$ the instantaneous upwards thrust and we suppose that the rate of decrease of mass is proportional to u, we obtain the following first–order system of ordinary differential equations

$$\begin{cases} x'(t) = y(t), \\ \\ y'(t) = -g + \dfrac{u(t)}{z(t)}, \\ \\ z'(t) = -Ku(t), \end{cases} \tag{2.39}$$

where $K > 0$ and g is the gravity acceleration. At time 0 we have the initial conditions

$$x(0) = x_0, \qquad y(0) = y_0, \qquad z(0) = z_0. \tag{2.40}$$

In addition we suppose that the thrust cannot exceed some fixed value, say $0 \leq u(t) \leq R$ for some $R > 0$. The vehicle will land softly at time $T \geq 0$ if

$$x(T) = 0, \qquad y(T) = 0.$$

The problem of soft landing is then to minimize the amount of fuel consumed from time 0 to time T, that is $z_0 - z(T)$. The problem actually includes two state constraints, namely

$$x(t) \geq 0 \quad \text{and} \quad z(t) \geq m_0,$$

where m_0 is the mass of the vehicle with empty fuel tanks. Hence the general formulation of the problem is

$$\inf\left\{ z_0 - z(T,u)\,dt : \begin{array}{c} u \in C_*([0,\infty),[0,R]) \text{ such that } \exists\, T \geq 0 \\ \text{with } (x(T,u),y(T,u),z(T,u)) \in C \\ \text{and } x(t,u) \geq 0 \;\forall\, t \in [0,T] \end{array} \right\}, \tag{2.41}$$

where of course $x(\cdot,u), y(\cdot,u), z(\cdot,u)$ are the solutions of (2.39)–(2.40) and

$$C = \{(0,0,z) \in \mathbb{R}^3 \;:\; z \geq m_0\}.$$

Observe that (2.39) gives

$$z'(t) = -K(y'(t)+g)z(t)$$

and then $z(t) = z_0 e^{-K\int_0^t (y'(s)+g)\,ds} = z_0 e^{-K(y(t)-y_0+gt)}$. So if $(x(T,u),y(T,u),z(T,u)) \in C$ we have that

$$z_0 - z(T,u) = z_0\left(1 - e^{-K(-y_0+gT)}\right).$$

Since the function $t \mapsto z_0\left(1 - e^{-K(-y_0+gt)}\right)$ is increasing, it can be proved that a control u is optimal for problem (2.39) if and only if u is optimal for the problem

$$\inf\left\{ T\,dt : \begin{array}{c} T = T(u),\; u \in C_*([0,\infty),[0,R]) \\ (x(T,u),y(T,u),z(T,u)) \in C \\ \text{and } x(t,u) \geq 0 \;\forall\, t \in [0,T] \end{array} \right\}, \tag{2.42}$$

where $x(\cdot,u), y(\cdot,u), z(\cdot,u)$ are the solutions of (2.39)–(2.40) and C as above. If we disregard condition $x(u,t) \geq 0$, the previous problem is a time optimal problem. Hence we will try to solve it by means of (a modification of) Theorem 2.2.16 (and Remark 2.2.17) and then verify condition $x(u,t) \geq 0$ a fortiori. So let $\{\bar{u},(\bar{x},\bar{y},\bar{z})\}$ be an optimal pair for problem (2.42) without condition $x(u,t) \geq 0$ and put $\bar{T} = T(\bar{u})$. It can be shown, with an argument similar to the one in Theorem 2.2.16, that the same assertion holds true in the case when C is a half–line, provided the co–state $\begin{pmatrix} \xi \\ \eta \\ \zeta \end{pmatrix}$ satisfies the following final condition

$$\begin{pmatrix} \xi(\bar{T}) \\ \eta(\bar{T}) \\ \zeta(\bar{T}) \end{pmatrix} = \begin{pmatrix} \bar{\xi} \\ \bar{\eta} \\ \bar{\zeta} \end{pmatrix},$$

where $\begin{pmatrix} \bar{\xi} \\ \bar{\eta} \\ \bar{\zeta} \end{pmatrix}$ is some element of the set $N_C(\bar{T}) = N_C((\bar{x}(\bar{T}),\bar{y}(\bar{T}),\bar{z}(\bar{T})))$, defined to be the set

$$\left\{ \begin{pmatrix} \bar{\xi} \\ \bar{\eta} \\ \bar{\zeta} \end{pmatrix} : \left|\begin{pmatrix} \bar{\xi} \\ \bar{\eta} \\ \bar{\zeta} \end{pmatrix}\right| = 1 \text{ and } \left\langle \begin{pmatrix} \bar{\xi} \\ \bar{\eta} \\ \bar{\zeta} \end{pmatrix}, \begin{pmatrix} x \\ y \\ z \end{pmatrix} - \begin{pmatrix} \bar{x}(\bar{T}) \\ \bar{y}(\bar{T}) \\ \bar{z}(\bar{T}) \end{pmatrix} \right\rangle \leq 0 \;\forall \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in C \right\}.$$

Now the adjoint system related to our problem is

$$\begin{cases} \xi'(t) = 0, \\ \eta'(t) = \xi(t), \\ \zeta'(t) = -\dfrac{u(t)}{z^2(t)}\eta(t), \end{cases} \qquad t \in (0, \bar{T})$$

and since $(\bar{x}(\bar{T}), \bar{y}(\bar{T}), \bar{z}(\bar{T})) = (0, 0, \bar{z})$ with $\bar{z} \geq m_0$, then $N_C(\bar{T})$ is either the set

$$\left\{ \begin{pmatrix} \bar{\xi} \\ \bar{\eta} \\ \bar{\zeta} \end{pmatrix} : \left| \begin{pmatrix} \bar{\xi} \\ \bar{\eta} \\ \bar{\zeta} \end{pmatrix} \right| = 1 \text{ and } \bar{\zeta} = 0 \right\} \qquad \text{if } \bar{\zeta} > m_0,$$

or

$$\left\{ \begin{pmatrix} \bar{\xi} \\ \bar{\eta} \\ \bar{\zeta} \end{pmatrix} : \left| \begin{pmatrix} \bar{\xi} \\ \bar{\eta} \\ \bar{\zeta} \end{pmatrix} \right| = 1 \text{ and } \bar{\zeta} \leq 0 \right\} \qquad \text{if } \bar{\zeta} = m_0.$$

In any case, $\zeta(\bar{T}) \leq 0$ and we get

$$\begin{cases} \xi(t) = \xi_0, \\ \eta(t) = \eta_0 - \xi_0 t, \\ \zeta'(t) = -\dfrac{\bar{u}(t)}{\bar{z}^2(t)}\eta(t), \\ \zeta(\bar{T}) \leq 0, \end{cases} \qquad t \in (0, \bar{T})$$

where ξ_0, η_0 are constants. Let us apply the Pontryagin Minimum Principle. We have

$$\bar{u}(t) \in \text{argmin}_{u \in [0,T]} \left\{ \xi_0 \bar{y}(t) + \left(\frac{u}{\bar{z}(t)} - g \right) \eta(t) - K u \zeta(t) \right\}$$

and then

$$\bar{u}(t) \in \text{argmin}_{u \in [0,T]} \left\{ u \left[\frac{\eta(t)}{\bar{z}(t)} - K\zeta(t) \right] \right\}.$$

In order to determine $\bar{u}$, let us prove that the map $\rho(t) := \frac{\eta(t)}{z(t)} - K\zeta(t)$ vanishes at most once on the interval $[0, \bar{T}]$. Indeed,

$$\rho'(t) = -\frac{\xi_0}{\bar{z}(t)} - \frac{K\eta(t)\bar{z}'(t)}{\bar{z}^2(t)} - K\zeta'(t) = -\frac{\xi_0}{\bar{z}(t)}.$$

There are several possibilities.

(a) $\xi_0 \neq 0$. Since $\bar{z}(\bar{T}) \geq m_0$ and $\bar{z}(t)$ is nonincreasing, we have that ρ is strictly increasing or decreasing, depending on the sign of ξ_0.

(b) $\xi_0 = 0$. Then $\rho(t) \equiv c$ and we must show that $c \neq 0$. There are two subalternatives.

(b1) $\eta_0 = 0$. This gives $\eta(t) \equiv 0$ and $\zeta(t) \equiv 1$, since the co–state is unitary. But then $\rho(t) \equiv -K$ is non–zero.

(b2) $\eta_0 \neq 0$. Since $\xi_0 = 0$, if $\rho \equiv 0$ then by Remark 2.2.17 we get $-g\eta_0 = 0$, which is a contradiction.

Since ρ can vanish at most once we obtain that there exists a time $t_0 \in [0, \bar{T}]$ such that ρ is strictly positive on $[0, t_0)$ and strictly negative on $(t_0, \bar{T}]$, or viceversa. In the interval where $\rho > 0$ we then have $\bar{u} \equiv 0$, while $\bar{u} \equiv R$ in the negative case. Notice that actually ρ can only change from positive to negative values (or remain strictly negative all time), since otherwise $\bar{u}$ would be 0 near the landing time, giving $\bar{y}(\bar{T}) \neq 0$. We can summarize the foregoing discussion saying that any optimal control $\bar{u}$ of the minimum time problem must be of the form

$$\bar{u}(t) = \begin{cases} 0 & \text{for } t \in [0, t_0) \\ R & \text{for } t \in [t_0, \bar{T}] \end{cases} \tag{2.43}$$

where either $t_0 = 0$ if $\rho < 0$ or $t_0 \in (0, T]$ the unique point where ρ vanishes. Let us now solve system (2.39) for a control of the form (2.43). In the full thrust interval $[t_0, \bar{T}]$ we have

$$\bar{z}(t) = z_0 - KR(t - t_0),$$

$$\bar{y}(t) = \bar{y}(t_0) - g(t - t_0) - \frac{1}{K} \log \frac{z_0 - KR(t - t_0)}{z_0}$$

and then

$$\begin{aligned} \bar{x}(t) &= \bar{x}(t_0) + \bar{y}(t_0)(t - t_0) - \frac{1}{2} g(t - t_0)^2 + \frac{t - t_0}{K} \\ &+ \frac{z_0 - KR(t - t_0)}{K^2 R} \log \frac{z_0 - KR(t - t_0)}{z_0} \end{aligned}$$

Since we are seeking for trajectories with $\bar{x}(\bar{T}) = \bar{y}(\bar{T}) = 0$, setting $s = \bar{T} - t_0$ we get

$$\phi(s) := gs + \frac{1}{K} \log \frac{z_0 - KRs}{z_0} = \bar{y}(t_0),$$

$$\psi(s) := -\bar{y}(t_0)s + \frac{1}{2} gs^2 - \frac{s}{K} - \frac{z_0 - KRs}{K^2 R} \log \frac{z_0 - KRs}{z_0} = \bar{x}(t_0)$$

and then, substituting $\bar{y}(t_0)$ in the last expression,

$$\psi(s) = -\frac{1}{2} gs^2 - \frac{s}{K} - \frac{z_0}{K^2 R} \log \frac{z_0 - KRs}{z_0} = \bar{x}(t_0)$$

Since the amount of fuel used up in the interval $[t_0, \bar{T}]$ is given by Rs, then

$$s \leq \frac{z_0 - m_0}{KR} < \frac{z_0}{KR}.$$

We assume that

$$\frac{R}{z_0} > g, \tag{2.44}$$

in order to offset the gravity acceleration as the engine is started at the beginning of the full thrust interval. This condition implies

$$\phi'(s) = g - \frac{R}{z_0 - KRs} < 0$$

so that ϕ is strictly decreasing in $0 \leq s < \frac{z_0}{KR}$. On the other hand,

$$\psi'(s) = -gs + \frac{z_0}{K(z_0 - KRs)} - \frac{1}{K},$$

$$\psi''(s) = -g + \frac{z_0 R}{(z_0 - KRs)^2} > -g + \frac{R}{z_0} > 0.$$

Since $\psi'(0) = 0$, then ψ is strictly increasing in $0 \leq s < \frac{z_0}{KR}$. In the initial interval $[0, t_0]$ of the trajectory, the rocket performs a free fall, so that we have

$$y(t) = -gt + y_0, \qquad x(t) = -\frac{1}{2}gt^2 + y_0 t + x_0. \tag{2.45}$$

The switching point t_0 and the terminal time $\bar{T}$ can be simultaneously determined by solving the system of two equations

$$\phi(\bar{T} - t_0) = -gt_0 + x_0, \qquad \psi(\bar{T} - t_0) = -\frac{1}{2}gt_0^2 + y_0 t_0 + x_0.$$

We can solve the system graphically as follows. Plot the curve $y = \phi(s)$ and $x = \psi(s)$ in the (x, y) plane for $0 \leq s < \frac{z_0}{KR}$, which is the *full thrust curve*. Then, starting from the initial position (x_0, y_0) draw the *free fall curve* given by (2.45) until it hits the full thrust curve. The value t_0 where this happens is the switching time; the value s_0 of the parameter s on the full thrust curve where the two curves intersect gives the final time $\bar{T} = s_0 + t_0$. After the intersection the trajectory of the rocket is given by $y(t) = \phi(\bar{T} - t)$ and $x(t) = \psi(\bar{T} - t)$, $t_0 \leq t \leq \bar{T}$. Observe that no initial positions below the full thrust curve can be steered to the origin.

2.3 Basic Problem in Dynamic Programming

Fix $T > 0$. For any $0 \leq s \leq T$ and $x \in \mathbb{R}^n$ consider the optimal control problem of minimization of the functional

$$J^{s,x}(u) = \int_s^T L(t, y_u^{s,x}(t), u(t))\; dt + \phi(y_u^{s,x}(T)) \tag{2.46}$$

among all the admissible controls $u(\cdot) \in C_*([s,T],U)$, where $y_u^{s,x}(t)$ is the solution of the controlled system

$$\begin{cases} y'(t) = f(t, y(t), u(t)) \\ \\ y(s) = x. \end{cases} \tag{2.47}$$

As before, we suppose $U \in \mathbb{R}^m$ closed, L and ϕ continuous and bounded from below functions, and f satisfying (2.4).

Remark 2.3.1 We have seen in Section 2.2 that the Pontryagin Maximum/Minimum Principle is a necessary condition to be satisfied by the optimal solution. Hence, if from this principle we obtain only one solution, then we determine the optimal trajectory and the optimal control; otherwise the Pontryagin Principle is not that useful. But by means of the Dynamic Programming approach (as we will see in this section) we are able to obtain a Necessary and Sufficient condition for optimality. Moreover, from the Pontryagin Maximum/Minimum Principle we get open loop controls, that is for each initial state we find $u = u(t)$, while from Dynamic Programming we get controls $u = u(t,x)$ such that we can solve the closed loop equation

$$\begin{cases} y'(t) = f(t, y(t), u(t, y(t))) \\ \\ y(s) = x, \end{cases}$$

and the related minimization problem (2.46) for any initial state. The control function $u = u(t,x)$ is called *closed loop control* or *feedback control.*

We define the value function v associated with (2.46)-(2.47) as $v : [0,T] \times \mathbb{R}^n \to \mathbb{R}$

$$v(s,x) = \min_{u(\cdot)} J^{s,x}(u). \tag{2.48}$$

The next Theorem is the Bellman Optimality Principle

Theorem 2.3.2 (Bellman's Optimality Principle) *Under the above assumptions on f, L, ϕ, the following assertions hold true.*
i) For any fixed $(s,x) \in [0,T] \times \mathbb{R}^n$ and any $u \in C_*([0,T],U)$

$$v(s,x) \leq \int_s^r L(t, y_u^{s,x}(t), u(t))\ dt + v(r, y_u^{s,x}(r)) \qquad \forall r \in [s,T]. \tag{2.49}$$

ii) $u^*(\cdot)$ is optimal if and only if

$$v(s,x) = \int_s^r L(t, y_{u^*}^{s,x}(t), u^*(t))\ dt + v(r, y_{u^*}^{s,x}(r)) \qquad \forall r \in [s,T]. \tag{2.50}$$

Proof– To prove statement i), consider an admissible strategy $u(\cdot)$ on $[s,T]$ and a strategy $\bar{u}(\cdot)$ on [r,T], where $r \in [s,T]$. Define a new strategy $\tilde{u}(\cdot)$ as follows

$$\tilde{u}(t) = \begin{cases} u(t) & \text{if } s \leq t \leq r \\ \\ \bar{u}(t) & \text{if } r < t \leq T \end{cases}$$

By definition of v we have

$$\begin{aligned} v(s,x) &\leq J^{s,x}(\tilde{u}) = \int_s^r L(t, y_u^{s,x}(t), u(t))\ dt + \int_r^T L(t, y_{\bar{u}}^{r, y_u^{s,x}(r)}(t), \bar{u}(t))\ dt + \phi(y_{\bar{u}}^{r,y_u^{s,x}(r)}(T)) \\ &= \int_s^r L(t, y_u^{s,x}(t), u(t))\ dt + J^{r, y_u^{s,x}(r)}(\bar{u}). \end{aligned}$$

Hence the first statement follows taking the minimum over $\bar{u}(\cdot)$ on both sides of the above equality.
In order to prove statement ii), first notice that if (2.50) holds, then u^* is the minimizer (take $r = T$). On the other hand if u^* is optimal then

$$\int_s^T L(t, y_{u^*}^{s,x}(t), u^*(t))\ dt + \phi(y_{u^*}^{s,x}(T)) = v(s,x),$$

which can be rewritten as

$$\int_r^T L(t, y_{u^*}^{s,x}(t), u^*(t))\ dt + \phi(y_{u^*}^{s,x}(T)) = v(s,x) - \int_s^r L(t, y_{u^*}^{s,x}(t), u^*(t))\ dt.$$

From inequality (2.49) and by the definition of v

$$v(s,x) - \int_s^r L(t, y_{u^*}^{s,x}(t), u^*(t))\ dt \leq v(r, y_{u^*}^{s,x}(r)) \leq \int_r^T L(t, y_{u^*}^{s,x}(t), u^*(t))\ dt + \phi(y_{u^*}^{s,x}(T)),$$

Hence (2.50) holds. ∎

Another way to state Bellman's Principle is to use the function V defined as follows

$$V(r) = \int_s^r L(t, y_u^{s,x}(t), u(t))\ dt + v(r, y_u^{s,x}(r)).$$

Theorem 2.3.3 *For any strategy $u(\cdot)$, $V(r)$ is non-decreasing, i.e. $V'(r) \geq 0$. If u^* is optimal then $V(r)$ is constant.*

Proof– Left as exercise.

Recall that the pre–Hamiltonian function $\mathcal{H} = \mathcal{H}(s,x,u,p)$ is given by

$$\mathcal{H}(s,x,u,p) = < p, f(s,x,u) > + L(s,x,u).$$

The *Hamiltonian function* is then defined by

$$H(t,x,p) := \min_{u \in U} \mathcal{H}(t,x,u,p). \tag{2.51}$$

The next Theorem can be seen as a link between Dynamic Programming Theory and Optimal Control Theory, since it shows that if v is regular enough, then ∇v is the co–state λ associated with the optimal trajectory. In what follows, we will denote by ∇v the gradient of v with respect to the x variable, while v_s will be the derivative of v with respect to s.

Theorem 2.3.4 *Assume v is everywhere differentiable. Then v is a solution of the Hamilton–Jacobi Cauchy problem*

$$\begin{cases} w_s(s,x) + H\left(s, x, \nabla w(s,x)\right) = 0, & \text{in } (0,T) \times \mathbb{R}^n \\ \\ w(T,x) = \phi(x). \end{cases} \tag{2.52}$$

Proof–It is easy to see that the terminal condition $v(T,x) = \phi(x)$ is fulfilled as a straightforward consequence of the definition of v. To show that v satisfies the Hamilton–Jacobi equation we evaluate the derivative of V. Taking $u(t) \equiv u$, $u \in U$ fixed, we have

$$0 \leq V'(r) = L(r, y_u^{s,x}(r), u) + v_s(r, y_u^{s,x}(r)) + < \nabla v(r, y_u^{s,x}(r)), f(r, y_u^{s,x}(r), u) >,$$

for any $r \in [s,T]$. In particular, for $r = s$ we obtain

$$0 \leq v_s(s,x) + L(s,x,u) + < \nabla v(s,x); f(s,x,u) > .$$

and hence

$$0 \leq v_s(s,x) + \min_{u \in U} \mathcal{H}(s,x,u,\nabla v(s,x)) = v_s(s,x) + H(s,x,\nabla v(s,x)).$$

On the other hand, if $u(t) = u^*(t)$, where u^* is optimal, this inequality is actually an equality. Indeed, V is constant and then

$$0 = v_s(s,x) + L(s,x,u^*(s)) + < \nabla v(s,x); f(s,x,u^*(s)) >,$$

which gives $v_s(s,x) + H(s,x,\nabla v(s,x)) \leq 0$. ■

Definition 2.3.5 *Equation*

$$0 = w_s(s,x) + H(s,x,\nabla w(s,x)),$$

is the Hamilton–Jacobi–Bellman Equation *or the* Dynamic Programming Equation*.*

Remark 2.3.6 If we are considering a maximization problem whose value function is defined as

$$v(s,x) = \max_{u(\cdot)} J^{s,x}(u),$$

then v is a solution of the Hamilton–Jacobi Cauchy problem

$$\begin{cases} w_s(s,x) + \max_{u \in U} \mathcal{H}(s,x,u,\nabla w(s,x)) = 0, & \text{in } (0,T) \times \mathbb{R}^n \\ \\ w(T,x) = \phi(x). \end{cases}$$

Let $Q_T = (0,T) \times \mathbb{R}^n$. We are going to state a sufficient condition for optimality. We begin by showing that any solution w of the Hamilton–Jacobi equation (2.52) is smaller than the value function v.

Theorem 2.3.7 *Let $w \in C^1(Q_T) \cap C(\overline{Q}_T)$ be a solution of (2.52). Then*

$$w \leq v.$$

Proof– Let $(s,x) \in Q_T$ be any initial state, $u(\cdot) \in C_*([s,T],U)$ any fixed strategy and $y(\cdot)$ the corresponding trajectory. Differentiating w with respect to s and recalling that w is a solution of (2.52), we get

$$\frac{d}{ds}w(t,y(t)) = w_s(t,y(t)) + \nabla w(t,y(t)) \cdot f(t,y(t),u(t))$$

$$= -H(t,y(t),\nabla w(t,y(t))) + \nabla w(t,y(t)) \cdot f(t,y(t),u(t))$$

$$\geq -L(t,y(t),u(t)).$$

Therefore,

$$\frac{d}{ds}w(t,y(t)) \geq -L(t,y(t),u(t)). \tag{2.53}$$

Recalling the terminal condition and integrating (2.53) from s to T we obtain

$$\phi(y(T)) - w(s,x) \geq -\int_s^T L(t,y(t),u(t))\ dt,$$

that is

$$J^{s,x}(u) \geq w(s,x).$$

The result now follows taking the minimum over $u(\cdot)$. ∎

Now we can state the sufficient condition for optimality, called *Verification Theorem.*

Theorem 2.3.8 *Let $w \in C^1(Q_T) \cap C(\overline{Q}_T)$ be a solution of (2.52) and let $(s,x) \in Q_T$. If there exists $u^*(\cdot) \in C_*([s,T],U)$ such that*

$$u^*(t) \in \text{argmin}\ \mathcal{H}(t,y^*(t),u,\nabla w(t,y^*(t))), \qquad \forall t \in [s,T], \tag{2.54}$$

where $y^*(\cdot) = y(\cdot,s,x,u^*)$, then $u^*(\cdot)$ is optimal and $w(s,x) = v(s,x)$.

Proof– Since y^* is the optimal trajectory associated with u^*, then from (2.54)

$$\frac{d}{dt}w(t,y^*(t)) = -L(t,y^*(t),u^*(t)).$$

Integrating from s to T we obtain

$$\phi(y^*(T)) - w(s,x) = -\int_s^T L(t,y^*(t),u^*(t))\ dt$$

and hence

$$w(s,x) = J^{s,x}(u^*) \geq v(s,x).$$

From the above inequality and from the previous Theorem it follows that $w(s,x) = v(s,x)$; therefore u^* is optimal. ∎

Remark 2.3.9 Suppose that for any $(s,x)\in Q_T$ we can find some

$$u^*(s,x)\in \text{argmin}\mathcal{H}(s,x,u,\nabla w(s,x)).$$

In general such a function can be very irregular. But if a selection of minimizers of $\mathcal{H}$ can be done in such a way that $u^*(\cdot,\cdot)$ is continuous, then the closed loop equation

$$\begin{cases} y'(t)=f(t,y(t),u^*(t,y(t))) \\ \\ y(s)=x. \end{cases}$$

automatically produces an optimal trajectory y^* of the mimimization problem (2.46)-(2.47). Moreover, the (open loop) control $\bar{u}^*(t):=u^*(t,y^*(t))$ is the optimal control whose related trajectory is y^*.

Definition 2.3.10 *The continuous control* u^* *in the previous remark is called* closed loop control *or* feedback.

We have established a three steps procedure in Dynamic Programming, that is:
1) Solve the Hamilton–Jacobi equation to get solutions $w\in C^1(Q_T)\cap C(\overline{Q}_T)$;
2) for every $(t,y)\in Q_T$ provide the feedback $u(t,y)\in \arg\min_u \mathcal{H}(t,y,u,\nabla w(t,y))$;
3) solve the closed loop equation

$$\begin{cases} y'(t)=f(t,y(t),u(t,y(t))) \\ \\ y(s)=x. \end{cases}$$

to get the optimal trajectories.

Now we apply this procedure to the Linear Quadratic Regulator problem

Example 2.3.11 This problem is known as the Linear Quadratic Regulator problem, that is the problem

$$\min_{u(\cdot)} J^{s,x}(u):=\min_{u(\cdot)}\left\{\int_s^T [M(t)y_u^{s,x}(t)\cdot y_u^{s,x}(t)+N(t)u(t)\cdot u(t)]\,dt+Dy_u^{s,x}(T)\cdot y_u^{s,x}(T)\right\},$$

where $y_u^{s,x}(\cdot)$ is governed by the controlled system

$$\begin{cases} y'(t)=A(t)y(t)+B(t)u(t) \\ \\ y(s)=x. \end{cases} \tag{2.55}$$

Here $y_u^{s,x}(\cdot)\in\mathbb{R}^n$, $U=\mathbb{R}^m$, $A(\cdot)\in M_{n\times n}$, $B(\cdot)\in M_{n\times m}$, $M(\cdot)\in M_{n\times n}$ is symmetric and positive semi–definite, $N(\cdot)\in M_{m\times m}$ is symmetric and positive definite and $D\in M_{n\times n}$ is symmetric and positive definite. All the matrices are continuous. The pre–Hamiltonian function associated with the problem is

$$\mathcal{H}(t,x,u,p)=M(t)x\cdot x+N(t)u\cdot u+p\cdot[A(t)x+B(t)u].$$

Minimizers of $\mathcal{H}$ in $\mathbb{R}^m$ (which exist because N is positive definite) must satisfy

$$0 = D_u\mathcal{H}(t,x,u,p) = 2N(t)u + B^*(t)p;$$

hence the minimizer is unique and given by

$$u^* = -\frac{1}{2}N^{-1}(t)B^*(t)p.$$

The Hamilton–Jacobi equation associated with our problem is

$$\begin{cases} w_s(t,x) + A(t)x \cdot \nabla w(t,x) + M(t)x \cdot x \\ \qquad\qquad -\dfrac{1}{4}B(t)N^{-1}(t)B^*(t)\nabla w(t,x) \cdot \nabla w(t,x) = 0, \qquad t \in (s,T) \\ w(T,x) = Dx \cdot x. \end{cases} \tag{2.56}$$

We will now apply the above three steps procedure:
Step 1): we will first show that the solution of (2.56) is of the form $w(t,x) = P(t)x \cdot x$, with $P(t) \in M_{m\times m}$ positive semi–definite and symmetric, which is reasonable because of the assumptions on the data. Substituting w as above into (2.56) we get

$$\begin{cases} P'(t)x \cdot x + 2A(t)x \cdot P(t)x + M(t)x \cdot x \\ \qquad\qquad -B(t)N^{-1}(t)B^*(t)P(t)x \cdot P(t)x = 0, \qquad t \in (s,T) \\ P(T)x \cdot x = Dx \cdot x. \end{cases}$$

Therefore P must be a solution of the so called *Riccati Equation*

$$\begin{cases} P'(t) + P(t)A(t) + A^*(t)P(t) + M(t) - P(t)B(t)N^{-1}(t)B^*(t)P(t) = 0, \quad t \in (s,T) \\ P(T) = D. \end{cases} \tag{2.57}$$

Now set $Q(t) := P(T-t)$. It is easy to see that P solves (2.57) if and only if Q solves

$$\begin{cases} Q'(t) = Q(t)\tilde{A}(t) + \tilde{A}^*(t)Q(t) + \tilde{M}(t) - Q(t)\tilde{B}(t)\tilde{N}^{-1}(t)\tilde{B}^*(t)Q(t) = 0, \quad t \in (0,T-s) \\ Q(0) = D, \end{cases} \tag{2.58}$$

where the notation $\tilde{K}(t)$ stands for $K(T-t)$. The previous one is a Cauchy problem with locally Lipschitz state function. Hence it admits a unique solution Q on some maximal interval $[0,\tau) \subset [0,T-s)$. It is easy to see that the solution is symmetric, i.e. $Q = Q^*$, since both Q and Q^* solve the same equation with the same initial condition. Moreover, Q is positive semi–definite. Indeed, let's go back to the matrix P. The previous argument shows the existence and uniqueness of a symmetric solution of the Riccati equation in the interval $(t_0, T] := (T - \tau, T]$, which is sufficient to have a feedback in this interval and to write the closed loop equation there. So fix any $x \in \mathbb{R}^n$, $t_1 \in (t_0, T)$ and consider the linear quadratic problem in the interval $[t_1, T]$ with initial condition given in t_1, $y(t_1) = x$ and cost functional $J^{t_1,x}(u)$. Let $u_1(\cdot)$ be the optimal control and $y_1(\cdot)$ the corresponding optimal trajectory. Then we have

$$
\begin{aligned}
&(P(t)y_1(t) \cdot y_1(t))' \\
&= P'(t)y_1(t) \cdot y_1(t) + 2P(t)y_1'(t) \cdot y_1(t) \\
&= \Big[- P(t)A(t) - A^*(t)P(t) - M(t) + P(t)B(t)N^{-1}(t)B^*(t)P(t)\Big] y_1(t) \cdot y_1(t) \\
&\quad + 2\Big[P(t)A(t) - P(t)B(t)N^{-1}(t)B^*(t)P(t)\Big] y_1(t) \cdot y_1(t) \\
&= -M(t)y_1(t) \cdot y_1(t) - N(t)u_1(t) \cdot u_1(t).
\end{aligned}
$$

Integrating in $[t_1, T]$ we then obtain

$$
P(t_1)x \cdot x = Dy_1(T) \cdot y_1(T) + \int_{t_1}^{T} \Big[M(t)y_1(t) \cdot y_1(t) + N(t)u_1(t) \cdot u_1(t)\Big]\, dt \geq 0,
$$

which gives the desired positivity property of P (and then of Q). It remains to show that actually $\tau \equiv T - s$, that is Q cannot "blow up". Indeed, on the space of symmetric and positive definite $n \times n$ matrices consider the following norm

$$
\|Q(t)\| := \max_{|x|=1} Q(t)x \cdot x,
$$

which is the one that gives the maximal eigenvalue of $Q(t)$ (and coincides with the usual operator norm on this space). Then, integrating (2.58) we get

$$
Q(t) - D = \int_0^t \Big[Q(r)\tilde{A}(r) + \tilde{A}^*(r)Q(r) + \tilde{M}(r) - Q(r)\tilde{B}(r)\tilde{N}^{-1}(r)\tilde{B}^*(r)Q(r)\Big]\, dr
$$

and then

$$
\begin{aligned}
Q(t)x \cdot x \quad &= \quad Dx \cdot x + \int_0^t \Big[Q(r)\tilde{A}(r)x \cdot x + \tilde{A}^*(r)Q(r)x \cdot x + \tilde{M}(r)x \cdot x\Big]\, dr \\
&\quad - \int_0^t \tilde{N}^{-1}(r)\tilde{B}^*(r)Q(r)x \cdot Q(r)\tilde{B}^*(r)x\, dr.
\end{aligned}
$$

Hence, by means of the Cauchy–Schwartz inequality

$$
|Q(s)\tilde{A}(s)x \cdot x| \leq \sqrt{Q(s)x \cdot x}\sqrt{Q(s)\tilde{A}(s)x \cdot \tilde{A}(s)x}
$$

we conclude that

$$
\|Q(t)\| \leq \|D\| + \int_0^t \Big[\|\tilde{M}(r)\| + 2\|Q(r)\|\|A(r)\|\Big] dr
$$

and the Gronwall inequality finally gives

$$\|Q(t)\| \le \left(\|D\| + \int_0^t \|\tilde{M}(r)\| \, dr \right) e^{2\int_0^t \|\tilde{A}(r)\| \, dr} \le C,$$

since the data are continuous. From the previous discussion we deduce that there exists a unique P solution of the Riccati Equation (2.57). Once we compute it explicitly, we can go on computing optimal trajectories and controls.
Step 2) We found previously that

$$u(t,y) = -\frac{1}{2} N^{-1}(t) B^*(t) P(t) y.$$

Step 3) To find the optimal trajectories solve the closed loop equation

$$\begin{cases} y'(t) = [A(t) - B(t)N^{-1}(t)B^*(t)P(t)]y(t), & t \in (s,T) \\ \\ y(s) = x. \end{cases}$$

Example 2.3.12 Suppose that $U = \mathbb{R}$ and consider the problem

$$\min \left\{ \int_0^1 u^2(t) \, dt + (x'(1))^2 \right\},$$

where x is the solution of

$$\begin{cases} x''(t) = u(t) \quad \text{in } (0,1) \\ \\ x(0) = x_0 \\ \\ x'(0) = y_0. \end{cases}$$

This problem can be rewritten as a Linear Quadratic Regulator problem as soon as we introduce another variable, namely $y(t) := x'(t)$. In this case we have

$$A(t) \equiv \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad B(t) \equiv \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad D = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix},$$

$M \equiv 0$ and $N \equiv 1$. Setting

$$Q(t) = \begin{pmatrix} \alpha(t) & \beta(t) \\ \\ \beta(t) & \gamma(t) \end{pmatrix},$$

it is easy to see that the Riccati equation for the matrix Q is given by

$$\begin{pmatrix} \alpha'(t) & \beta'(t) \\ \\ \beta'(t) & \gamma'(t) \end{pmatrix} = \begin{pmatrix} -\beta^2(t) & \alpha(t) - \beta(t)\gamma(t) \\ \\ \alpha(t) - \beta(t)\gamma(t) & 2\beta(t) - \gamma^2(t) \end{pmatrix},$$

$$\begin{pmatrix} \alpha(0) & \beta(0) \\ \\ \beta(0) & \gamma(0) \end{pmatrix} = D = \begin{pmatrix} 0 & 0 \\ \\ 0 & 1 \end{pmatrix}.$$

Hence the solution is

$$Q(t) = \begin{pmatrix} 0 & 0 \\ \\ 0 & \dfrac{1}{1+t} \end{pmatrix},$$

which in turn gives

$$P(t) = \begin{pmatrix} 0 & 0 \\ \\ 0 & \dfrac{1}{2-t} \end{pmatrix}.$$

Exercise 2.3.13 Complete the previous example, finding out optimal trajectories and controls.

Remark 2.3.14 In this section we found a link between the optimal control problem of minimizing

$$J^{s,x}(u) = \int_s^T L(t, y_u^{s,x}(t), u(t))\ dt + \phi(y_u^{s,x}(T))$$

on the admissible control $u(\cdot) \in C_*([s,T], U)$, where $y_u^{s,x}(t)$ is the solution of the controlled system

$$\begin{cases} y'(t) = f(t, y(t), u(t)) \\ \\ y(s) = x, \end{cases}$$

and the Hamilton–Jacobi Cauchy problem

$$\begin{cases} w_s(s,x) + \min\limits_{u \in U} \{\nabla w(t, y(t)) \cdot f(t, y(t), u) + L(t, y(t), u(t))\} = 0, & \text{in } (0,T) \times \mathbb{R}^n \\ \\ w(T, x) = \phi(x). \end{cases}$$

But if we consider the minimization of the functional

$$J^{s,x}(u) = \int_0^s L(t, y_u^{s,x}(t), u(t))\ dt + \phi(y_u^{s,x}(0))$$

on the admissible control $u(\cdot) \in C_*([0,s], U)$, where $y_u^{s,x}(t)$ is the solution of the controlled system

$$\begin{cases} y'(t) = f(t, y(t), u(t)) \\ \\ y(s) = x, \end{cases}$$

all the previous results hold true, provided we substitute the above Hamilton–Jacobi Cauchy problem with the following one

$$\begin{cases} w_s(s,x) + \max\limits_{u \in U} \{\nabla w(t, y(t)) \cdot f(t, y(t), u) - L(t, y(t), u(t))\} = 0, & \text{in } (0,T) \times \mathbb{R}^n \\ \\ w(0, x) = \phi(x). \end{cases}$$

Observe that in the case $f(t,x,u) = u$ and $U = \mathbb{R}^n$, as in Calculus of Variations theory,

$$\max_{u \in U} \{\nabla w(t, y(t)) \cdot f(t, y(t), u) - L(t, y(t), u(t))\}$$

is exactly the Legendre–Fenchel transform of L that we introduced in Chapter 1.

2.4 The Method of Characteristics

For the study of the Hamilton–Jacobi Cauchy problem (2.52) it is interesting to consider the method of characteristics, which is a classical approach to find regular solutions to first order partial differential equations. As we show below, the method of characteristics provides a local smooth solution of the equation, if the data are smooth, but at the same time shows that in general no global regular solution exists.

Let us consider the problem

$$\begin{cases} w_t(t,x) + H(t,x,\nabla w(t,x)) = 0, & (t,x) \in (0,T) \times \mathbb{R}^n, \\ w(0,x) = \phi(x), & x \in \mathbb{R}^n, \end{cases} \tag{2.59}$$

with H and ϕ of class C^2.

Let us suppose to have a solution $w \in C^2([0,T] \times \mathbb{R}^n)$ of the above problem. Given $z \in \mathbb{R}^n$, we call *characteristic curve* associated with w starting from z the curve $X(\cdot; z)$ which solves

$$\dot{X} = H_p(t, X, \nabla w(t,X)), \qquad X(0) = z.$$

Moreover, if we set

$$U(t;z) = w(t, X(t;z)), \qquad P(t;z) = \nabla w(t, X(t;z)).$$

we find that

$$\dot{U} = w_t(t,X) + \nabla w(t,X) \cdot \dot{X} = -H(t,X,P) + P \cdot H_p(t,X,P),$$

$$\begin{aligned} \dot{P} &= \nabla w_t(t,X) + \nabla^2 w(t,X) H_p(t,X,\nabla w(t,X)) \\ &= \nabla[w_t(t,X) + H(t,X,\nabla w(t,X))] - H_x(t,X,\nabla w(t,X)) \\ &= -H_x(t,X,P). \end{aligned}$$

Therefore, the pair X, P solves the ordinary differential problem

$$\begin{cases} \dot{X} = H_p(t,X,P) \\ \dot{P} = -H_x(t,X,P) \end{cases} \qquad \begin{cases} X(0) = z \\ P(0) = \nabla \phi(z), \end{cases} \tag{2.60}$$

while U satisfies

$$\dot{U} = -H(t,X,P) + P \cdot H_p(t,X,P), \qquad U(0;z) = \phi(z). \tag{2.61}$$

This shows that X, P and U are uniquely determined by the initial value ϕ. The above arguments suggest that one can obtain a solution of the Hamilton–Jacobi equation by solving (2.60), as long as the map $z \to X(t;z)$ is invertible. Indeed we have the following result.

Theorem 2.4.1 *Given $z \in \mathbb{R}^n$, let $X(t;z)$, $P(t;z)$ denote the solution of problem* (2.60) *and let $U(t;z)$ be defined by* (2.61). *Suppose that there exists $T^* > 0$ such that*

(i) *the maximal interval of existence of the solution to* (2.60) *contains $[0, T^*[$ for all $z \in \mathbb{R}^n$;*

(ii) *the map $z \to X(t;z)$ is invertible with C^1 inverse $x \to Z(t;x)$ for all $t \in [0, T^*[$.*

Then there exists a unique solution $w \in C^2([0, T^*[\times \mathbb{R}^n)$ of problem (2.59), *which is given by*

$$w(t,x) = U(t; Z(t;x)), \qquad (t,x) \in [0, T^*[\times \mathbb{R}^n.$$

Proof — For simplicity, in what follows we will often write P, Z, etc. instead of $P(t; Z(t;x))$, $Z = Z(t;x)$ respectively.
From the definition it is clear that w is of class C^1. Let us compute its derivatives. Observe first that by (2.60) X_z is a solution of the equation

$$\dot{X}_z = H_{px}(t,X,P)X_z + H_{pp}(t,X,P)P_z, \quad X_z(0) = I$$

and U_z solves the equation[2]

$$\dot{U}_z = -H_x(t,X,P)X_z + P[H_{px}(t,X,P)X_z + H_{pp}(t,X,P)P_z],$$

with initial condition $U_z(0) = \nabla\phi(z)$. Hence it is easy to see that $U_z = (PX_z)$, since (PX_z) and U_z both solve the same equation and $(PX_z)(0) = U_z(0) = \nabla\phi(z)$. This implies, by the definition of Z,

$$\nabla w(t,x) = U_z(t;Z)Z_x = (P(t;Z)X_z)Z_x = P(t; Z(t;x)). \tag{2.62}$$

We also find

$$\begin{aligned} w_t(t,x) &= U_t(t;Z(t;x)) + U_z(t;Z(t;x)) \cdot Z_t(t;x) \\ &= [PH_p(t,X,P) - H(t,X,P)] + (PX_z) \cdot Z_t. \end{aligned} \tag{2.63}$$

Taking into account that $X(t; Z(t;x)) \equiv x$ we obtain

$$X_t(t;Z) + X_z(t;Z)Z_t = 0. \tag{2.64}$$

Therefore, by (2.62), (2.63) and (2.64)

$$\begin{aligned} & w_t(t,x) + H(t,x,\nabla w(t,x)) = \\ &= [P \cdot H_p(t,X,P) - H(t,X,P)] + (PX_z) \cdot Z_t + H(t,X,P) = \\ &= P \cdot H_p(t,X,P) + (PX_z) \cdot Z_t = \\ &= P \cdot [X_t + X_z Z_t] = 0. \end{aligned}$$

[2]In what follows we will use the standard notation bA to indicate the left product of a vector $b \in \mathbb{R}^n$ with a $N \times N$ matrix A, while $a \cdot b$ will denote the scalar product between $a, b \in \mathbb{R}^n$.

This equality implies that $w \in C^2$ and satisfies problem (2.59) (the fact that the initial condition is satisfied is a direct consequence of the definition). Uniqueness follows from the assumption that the map $z \to X(t;z)$ is invertible with C^1 inverse. ■

Example 2.4.2 Let us apply the previous results to find the (unique) solution of problem

$$\begin{cases} w_t(t,x) + \dfrac{1}{2} w_x^2(t,x) = \dfrac{a}{2}x^2 & \text{for } (t,x) \in \mathbb{R}^+ \times \mathbb{R}, \\ \\ w(0,x) = \lambda x^2 & \text{for all } x \in \mathbb{R}. \end{cases}$$

where a, λ are positive constants. Since the Hamiltonian is given by $H(t,x,p) = \frac{1}{2}p^2 - \frac{a}{2}x^2$, the equations (2.60) become

$$\begin{cases} \dot{X} = P \\ \\ \dot{P} = aX \end{cases} \qquad \begin{cases} X(0) = z \\ \\ P(0) = 2\lambda z. \end{cases} \tag{2.65}$$

Solving these equations is equivalent to solve system

$$\begin{cases} \ddot{X} = aX \\ \\ X(0) = z \\ \\ \dot{X}(0) = 2\lambda z. \end{cases} \tag{2.66}$$

So $X(t;z) = c_1 e^{t\sqrt{a}} + c_2 e^{-t\sqrt{a}}$, where $c_1 = \frac{\sqrt{a}+2\lambda}{2\sqrt{a}}z$ and $c_2 = \frac{\sqrt{a}-2\lambda}{2\sqrt{a}}z$. Now set $\phi(t) = \frac{\sqrt{a}+2\lambda}{2\sqrt{a}}e^{t\sqrt{a}} + \frac{\sqrt{a}-2\lambda}{2\sqrt{a}}e^{-t\sqrt{a}}$. Since $\lambda > 0$, $\phi(t) \neq 0$ for all $t \in \mathbb{R}^+$. Hence $z \to X(t;z)$ is invertible for all t and the inverse map is $Z(t,x) = \frac{x}{\phi(t)}$. Applying Theorem 2.4.1 we deduce that the solution of our problem exists for all $t \in \mathbb{R}^+$. Now,

$$P(t,z) = \dot{X}(t;z) = \dot{\phi}(t)z = \left(\frac{\sqrt{a}+2\lambda}{2} e^{t\sqrt{a}} - \frac{\sqrt{a}-2\lambda}{2} e^{-t\sqrt{a}} \right) z$$

and

$$\begin{aligned} U(t;z) &= \lambda z^2 + \int_0^t [|\dot{\phi}(s)z|^2 - \frac{1}{2}|\dot{\phi}(s)z|^2 + \frac{a}{2}|\phi(s)z|^2]\, ds \\ &= \lambda z^2 + \frac{z^2}{2}\int_0^t [\dot{\phi}(s)^2 + a\phi^2(s)]\, ds \\ &= z^2 \left(\frac{(\sqrt{a}+2\lambda)^2}{8\sqrt{a}} e^{2t\sqrt{a}} - \frac{(\sqrt{a}-2\lambda)^2}{8\sqrt{a}} e^{-2t\sqrt{a}} \right). \end{aligned}$$

Therefore,

$$\begin{aligned} w(t,x) &= U(t,Z(t;x)) = \frac{\lambda x^2}{\phi^2(t)} + \frac{x^2}{2\phi^2(t)}\int_0^t [\dot\phi(s)^2 + a\phi^2(s)]\,ds \\ &= \frac{\sqrt{a}x^2}{2}\frac{(\sqrt{a}+2\lambda)e^{t\sqrt{a}} - (\sqrt{a}-2\lambda)e^{-t\sqrt{a}}}{(\sqrt{a}+2\lambda)e^{t\sqrt{a}} + (\sqrt{a}-2\lambda)e^{-t\sqrt{a}}}. \end{aligned}$$

Example 2.4.3 We want to apply the Method of Characteristics to solve the following partial differential equation

$$\begin{cases} w_t(t,x) + \dfrac{1}{2}w_x(t,x)^2 = 0, & (t,x) \in (0,T)\times \mathbb{R} \\ w(0,x) = \frac{1}{2}x^2 \end{cases} \tag{2.67}$$

Here the Hamiltonian function is $H = H(p) = \frac{1}{2}p^2$ and the associated ordinary system is

$$\begin{cases} \dot X(s) = DH(P(s)) = P(s) & X(0) = z \\ \dot U(s) = P(s)DH(P(s)) - H(P(s)) = \frac{1}{2}P(s)^2 & U(0) = \phi(z) = \frac{1}{2}z^2 \\ \dot P(s) = 0 & P(0) = D\phi(z) = z, \end{cases}$$

whose solution is

$$\begin{cases} X(s,z) = (1+s)z \\ U(s,z) = \frac{1}{2}z^2(1+s) \\ P(s,z) = z. \end{cases}$$

Deducing z from the expression $x = X(t,z) = (1+t)z$ we get $z = Z(t,x) = \frac{x}{1+t}$ and evaluating U at $(t, Z(t,x))$ we obtain the solution w of (2.67)

$$w(t,x) = U(t,Z(t,x)) = \frac{1}{2}\frac{x^2}{1+t}.$$

Since the Hamiltonian H is a superlinear and strictly convex function (with respect to p), the system (2.67) can be associated with a problem of Calculus of Variations, whose Lagrangian is the Legendre–Fenchel transform of H (see Remark 2.3.14). In our case the Lagrangian is given by $L(t,x,q) = \frac{1}{2}q^2$ and then the problem associated with (2.67) is the following: for any fixed $(s,x) \in (0,T)\times\mathbb{R}$ minimize

$$J(\xi) = \left\{\int_0^s \frac{1}{2}\xi'(t)^2\,dt + \frac{1}{2}\xi(0)^2 : \xi \in A\right\},$$

where $A = \{\xi \in C^1_*([0,s];\mathbb{R}^n),\ \xi(s) = x\}$, or equivalently

$$J^{s,x}(u) = \left\{\int_0^s \frac{1}{2}y'^2(t)\,dt + \frac{1}{2}y^2(0)\right\},$$

on the admissible controls $u(\cdot) \in C_*([0,s], U)$, where $y(\cdot)$ is the solution of the controlled system

$$\begin{cases} y'(t) = u(t), & t \in (0,s) \\ \\ y(s) = x. \end{cases} \tag{2.68}$$

In this context, the pre–Hamiltonian function is $\mathcal{H}(s,x,u,p) = \frac{1}{2}u^2 + pu$. Hence $u^*(s,x) = \nabla w(s,x) = \frac{x}{1+s}$ is a continuous selection of $\mathrm{argmin}\mathcal{H}(s,x,u,\nabla w(s,x))$. Then, the closed loop equation related to our problem is

$$\begin{cases} y'(t) = \dfrac{x}{1+s}, & t \in (0,s) \\ \\ y(s) = x, \end{cases}$$

whose solution is

$$y^*(t) = \frac{1+t}{1+s}x.$$

Remark 2.4.4 In general, the procedure of making z explicit in the expression $x = X(t,z)$ can be applied only locally, that is for a short amount of time. For example, if we change the initial condition in (2.67) by

$$w(0,x) = -\frac{1}{2}x^2,$$

then applying the Method of Characteristics we have

$$Z(t,x) = \frac{x}{1-t},$$

hence the function w

$$w(t,x) = \frac{1}{2}\frac{x^2}{t-1}$$

is defined only for $t \in [0,1)$.

Remark 2.4.5 The optimization problem of minimizing

$$J(\xi) = \left\{ \int_0^s \frac{1}{2}\xi'^2(t)\, dt + \frac{1}{2}\xi(0)^2 | \xi \in A \right\},$$

where $A = \{\xi \in C^1_*([0,s]; \mathbb{R}^n) \mid \xi(s) = x\}$, can be directly solved using the Calculus of Variations method instead of using the Hamiltonian formulation. Since the Lagrangian depends on the q variable only, the optimal arcs are straight lines; hence we can write

$$\xi^*(t) = at + b, \tag{2.69}$$

where a and b will be determined exploiting the initial condition and the transversality condition. In fact, evaluating (2.69) at $t = s$, we have $\xi^*(s) = x = as + b$. Hence $b = x - as$ and substituting b in (2.69) we get $\xi^*(t) = a(t-s) + x$. On the other hand, since

$$\frac{\partial L}{\partial q}(\xi^{*\prime}(0)) = \xi^{*\prime}(0) = a \qquad \text{and} \qquad D\phi(0) = \xi^*(0) = -as + x,$$

the transversality condition $\dfrac{\partial L}{\partial q}(\xi^{*\prime}(0)) = D\phi(0)$ yields

$$a = -as + x, \qquad \text{that is} \qquad a = \frac{x}{1+s}.$$

Substituting a and b into (2.69) we obtain

$$\xi^*(t) = \frac{1+t}{1+s}x,$$

as in the previous example.

Exercise 2.4.6 *Solve the following problem by means of the Method of Characteristics.*

$$\begin{cases} w_t(t,x) + b \cdot \nabla w(t,x) = 0 & \text{for } (t,x) \in \mathbb{R}^+ \times \mathbb{R}^n, \\ w(0,x) = w_0(x) & \text{for all } x \in \mathbb{R}^n, \end{cases}$$

where $b \in \mathbb{R}^n$ is a fixed vector and w_0 is arbitrary.

Exercise 2.4.7 *Find the unique regular solution and the related maximal existence time T of problem*

$$\begin{cases} w_t(t,x) + \frac{1}{2}(w_x(t,x)^2 + x^2) = x & \text{for } x \in \mathbb{R},\ t > 0, \\ w(0,x) = x & \text{for all } x \in \mathbb{R}. \end{cases}$$

Chapter 3

Appendix A

3.1 The Legendre Transform

The *Legendre transform* (sometimes also called *Fenchel transform* or *convex conjugate*) is a classical topic of convex analysis. A comprehensive treatment of this transform can be found in many textbooks; here we give, for the convenience of the reader, a short and self–contained exposition of its properties in the C^1-case which is enough for the purposes of these notes.

Definition 3.1.1 *Let* $L : \mathbb{R}^n \to \mathbb{R}$ *be a* $C^1(\mathbb{R}^n, \mathbb{R})$ *and convex function which satisfies*

$$\lim_{|q|\to\infty} \frac{L(q)}{|q|} = +\infty, \tag{3.1}$$

i.e. L *is superlinear. The* Legendre transform *of* L *is the function*

$$L^*(p) = \sup_{q\in\mathbb{R}^n} [p \cdot q - L(q)], \qquad p \in \mathbb{R}^n. \tag{3.2}$$

Example 3.1.2 Define $L : \mathbb{R}^n \to \mathbb{R}$ by $L(q) = \frac{1}{2}|q|^2$.
For all $p \in \mathbb{R}^n$ the map $q \mapsto p \cdot q - L(q)$ is continuously differentiable at all points and goes to $-\infty$ as $|q| \to \infty$; thus it attains its maximum at a point q where its differential $p - DL(q)$ vanishes. This means that $p - q = 0$ and hence $L^*(p) = \frac{1}{2}|p|^2$.

Example 3.1.3 Given any symmetric, positive definite matrix A of order n, define $L : \mathbb{R}^n \to \mathbb{R}$ by $L(q) = \frac{1}{2}Aq \cdot q$.
As in the previous example, for all $p \in \mathbb{R}^n$ the map $q \mapsto p\cdot q - L(q)$ achieves its maximum at a point q where its differential $p - DL(q) = p - Aq$ vanishes. Since A is invertible, we get that $q = A^{-1}p$. Hence $L^*(p) = \frac{1}{2}A^{-1}p \cdot p$.

Exercise 3.1.4 *Calculate the Legendre transform of the following maps* L *defined over* $\mathbb{R}^n$*:*

- $L(q) = \frac{1}{2}Aq \cdot q + b \cdot q$, *where* A *is a positive definite matrix of order* n *and* $b \in \mathbb{R}^n$ *is a fixed vector;*

- $L(q) = \frac{1}{s}|q|^s$, *where* $s > 1$.

Theorem 3.1.5 *Let L be as in Definition 3.1.1. Then*

(a) *for every p there exists a point q_p where the supremum in* (3.2) *is attained. In addition, for every bounded set $C \subset \mathbb{R}^n$ there exists $R > 0$ such that $|q_p| < R$ for all $p \in C$.*

(b) *the function L^* is convex and superlinear.*

(c) *$L^{**} = L$; moreover, $p = DL(q)$ if and only if $L^*(p) + L(q) = p \cdot q$.*

Proof — **(a)** The claimed properties are a straightforward consequence of the convexity of L and of assumption (3.1). Notice that for every $p \in \mathbb{R}^n$ and q_p such that $L^*(p) = p \cdot q_p - L(q_p)$ we must have $DL(q_p) = p$, since $F(q) = p \cdot q - L(q)$ is a regular function which attains its maximum in q_p.
(b) Take any $p_1, p_2 \in \mathbb{R}^n$ and $t \in [0,1]$ and let q_t be a point such that

$$L^*(tp_1 + (1-t)p_2) = (tp_1 + (1-t)p_2) \cdot q_t - L(q_t).$$

Since $L^*(p_i) \geq p_i \cdot q_t - L(q_t)$ for $i = 1, 2$, we easily conclude that $L^*(tp_1 + (1-t)p_2) \leq tL^*(p_1) + (1-t)L^*(p_2)$, i.e. L^* is convex. In addition, for all $M > 0$ and $p \in \mathbb{R}^n$ we have

$$L^*(p) \geq M\frac{p}{|p|} \cdot p - L\left(M\frac{p}{|p|}\right) \geq M|p| - \max_{|q|=M} L(q),$$

and so, for all $M > 0$ and any sequence $\{p_k\}$ such that $p_k \to \infty$ as $k \to \infty$

$$\frac{L^*(p_k)}{|p_k|} \geq \frac{M}{2} \qquad \text{for } k \text{ sufficiently large.}$$

Since M is arbitrary, L^* turns out to be superlinear.
(c) By definition we have that $L(q) \geq q \cdot p - L^*(p)$ for all $q, p \in \mathbb{R}$, which implies that $L \geq L^{**}$. To prove the converse inequality, fix any $\bar{q} \in \mathbb{R}^n$ and take $\bar{v} = DL(\bar{q})$. Let $q(\bar{v})$ be such that $L^*(\bar{v}) = q(\bar{v}) \cdot \bar{v} - L(q(\bar{v}))$. Since $\bar{v} = DL(\bar{q})$, we have $L(q(\bar{v})) - L(\bar{q}) - \bar{v} \cdot (q(\bar{v}) - \bar{q}) \geq 0$ (by the convexity of L) and therefore

$$L(\bar{q}) \leq L(q(\bar{v})) - \bar{v} \cdot (q(\bar{v}) - \bar{q}) = -L^*(\bar{v}) + \bar{q} \cdot \bar{v},$$

which implies that $L(\bar{q}) \leq L^{**}(\bar{q})$. The second assertion is now a simple consequence of the above proof. ∎

Theorem 3.1.6 *Let L be a C^2 and strictly convex function, satisfying (3.1). Then*

(a) *for every p there exists a unique point q_p where the supremum in* (3.2) *is attained.*

(b) *L^* is a strictly convex and C^2 function.*

Proof — **(a)** Existence is already proved. Uniqueness is a consequence of the following facts:

- for every $p \in \mathbb{R}^n$ and q_p such that $L^*(p) = p \cdot q_p - L(q_p)$ we have $DL(q_p) = p$;
- $D_q L$ is injective, since, for every $q_1 \neq q_2$ in $\mathbb{R}^n$

$$(DL(q_1) - DL(q_2)) \cdot (q_1 - q_2) = \int_0^1 D^2L(tq_1 + (1-t)q_2)(q_1 - q_2)dt \cdot (q_1 - q_2) > 0.$$

(by assumption L is a strictly convex function)
Now, if there exist p and two points $q_1 \neq q_2$ in $\mathbb{R}^n$ such that $L^*(p) = p \cdot q_1 - L(q_1) = p \cdot q_2 - L(q_2)$, then $p = DL(q_1) = DL(q_2)$, against the injectivity of DL.
(b) By the strict convexity of L and the fact that DL is an injective and surjective map (see **(a)**), we have that DL is actually a global diffeomorphism of class C^1. Furthermore, for any $p \in \mathbb{R}^n$ $L^*(p) = p \cdot q_p - L(q_p)$, where $DL(q_p) = p$, i.e. $q_p = DL^{-1}(p)$. Hence $L^*(p) = p \cdot DL^{-1}(p) - L(DL^{-1}(p))$ is at least a C^1 function. Moreover, from the proof **(c)** of Theorem 3.1.5 we get that $L^{**} = L$ and that $\bar{p} = DL(\bar{q})$ if and only if $L(\bar{q}) + L^*(\bar{p}) = \bar{p} \cdot \bar{q}$. This means that any $\bar{p} \in \mathbb{R}^n$ maximizes $G(p) = \bar{q} \cdot p - L^*(p)$, where $\bar{q}$ is such that $\bar{p} = DL(\bar{q})$. Since L^* is a C^1 function, then $\bar{p}$ must satisfy $\bar{q} = DL^*(\bar{p}) = DL^{-1}(\bar{p})$. This means that DL^* coincides with the inverse of DL and hence it is itself a C^1 function. To conclude the proof, we have to show that L^* is strictly convex. But this follows from the fact that $DL^* = DL^{-1}$ and hence $D^2L^*(p) = \left[D^2L\left((DL)^{-1}(p)\right)\right]^{-1}$. ∎

Now let $\Omega \subset \mathbb{R}^n$ be an open set, $a, b \in \mathbb{R}$, $a < b$, and let $L \in C([a,b] \times \Omega \times \mathbb{R}^n)$. Suppose that L is convex in the third argument and that it satisfies

$$\lim_{|q| \to \infty} \inf_{(t,x) \in [a,b] \times \Omega} \frac{L(t,x,q)}{|q|} = +\infty. \tag{3.3}$$

Under these hypotheses we can define the Legendre transform of L with respect to the third group of variables as follows

$$H(t,x,p) = \sup_{q \in \mathbb{R}^n} \left[p \cdot q - L(t,x,q)\right], \qquad (t,x,q) \in [a,b] \times \Omega \times \mathbb{R}^n. \tag{3.4}$$

Theorem 3.1.7

(a) *For every $(t,x,p) \in [a,b] \times \Omega \times \mathbb{R}^n$ there exists at least one point $q(t,x,p)$ where the supremum in (3.4) is attained. In addition, for every bounded set $C \subset \mathbb{R}^n$ there exists $R > 0$ such that every q associated with $(t,x,p) \in [a,b] \times \Omega \times C$ satisfies $|q| < R$.*

(b) *The transformed function H is continuous, convex in the third argument and satisfies*

$$\lim_{|p| \to \infty} \inf_{(t,x) \in [a,b] \times \Omega} \frac{H(t,x,p)}{|p|} = +\infty.$$

Proof — Without loss of generality we can suppose that L does not depend on t, i.e. $L(t, x, q) = L(x, q)$.
(a) It is a direct consequence of the properties of L.
(b) Convexity and superlinearity are proved as for the case without x–dependence. To prove continuity, let us consider a sequence $(x_m, p_m) \subset \Omega \times \mathbb{R}^n$ converging to some $(\bar{x}, \bar{p})$. We have

$$H(x_m, p_m) = p_m \cdot q_m - L(x_m, q_m), \qquad H(\bar{x}, \bar{p}) = \bar{p} \cdot \bar{q} - L(\bar{x}, \bar{q})$$

for suitable $q_m, \bar{q}$. Then

$$H(x_m, p_m) \geq p_m \cdot \bar{q} - L(x_m, \bar{q}).$$

Letting $m \to \infty$ we obtain that $\liminf_{m\to\infty} H(x_m, p_m) \geq H(\bar{x}, \bar{p})$. To prove the converse inequality, let us choose a subsequence (x_{m_k}, p_{m_k}) such that $\lim_{k\to\infty} H(x_{m_k}, p_{m_k}) = \limsup_{m\to\infty} H(x_m, p_m)$. Since, by part (a), $\{q_m\}$ is bounded, we can assume that the subsequence q_{m_k} converges to some value q^*. Then we have

$$\begin{aligned} H(\bar{x}, \bar{p}) &\geq \bar{p} \cdot q^* - L(\bar{x}, q^*) = \lim_{k\to\infty} p_{m_k} \cdot q_{m_k} - L(x_{m_k}, q_{m_k}) \\ &= \limsup_{m\to\infty} H(x_m, p_m). \end{aligned}$$

This proves the continuity of H. ∎

To conclude we consider the case when L is strictly convex and smooth.

Corollary 3.1.8 *Let $L \in C^2([a, b] \times \Omega \times \mathbb{R}^n)$. Assume that L satisfies* (3.3) *and*

$$\frac{\partial^2 L}{\partial q^2}(t, x, q) \quad \text{is positive definite for all } (t, x, q) \in [a, b] \times \Omega \times \mathbb{R}^n. \tag{3.5}$$

Then H belongs to $C^2([a, b] \times \Omega \times \mathbb{R}^n)$. Moreover, if we denote by $q(t, x, p)$ the unique value of q at which the infimum in (3.4) *is attained, we have*

$$\frac{\partial H}{\partial p}(t, x, p) = q(t, x, p), \tag{3.6}$$

$$\frac{\partial H}{\partial x}(t, x, p) = -\frac{\partial L}{\partial x}(x, q(t, x, p)) \tag{3.7}$$

$$\frac{\partial H}{\partial t}(t, x, p) = -\frac{\partial L}{\partial t}(x, q(t, x, p)) \tag{3.8}$$

$$\frac{\partial^2 H}{\partial p^2}(t, x, p) = \left[\frac{\partial^2 L}{\partial q^2}(x, q(t, x, p))\right]^{-1} \tag{3.9}$$

In addition,

$$q = \frac{\partial H}{\partial p}(t, x, p) \qquad \text{if and only if} \qquad p = \frac{\partial L}{\partial q}(t, x, q) \tag{3.10}$$

Proof —As before, consider the case $L(t,x,q) = L(x,q)$.
For any $x \in \Omega$, Theorem 3.1.6 implies that $H(x,\cdot) \in C^2(\mathbb{R}^n)$ and that (3.10) holds. Since $q(x,p)$ satisfies

$$p - \frac{\partial L}{\partial q}(x, q(x,p)) = 0 \tag{3.11}$$

we see that (3.6) follows from (3.10) and that (3.9) can be obtained as in Theorem 3.1.6-(b). In addition, from the Implicit Function Theorem and from assumption (3.5) we get that $q(x,p) \in C^1(\Omega \times \mathbb{R}^n)$.
Moreover, recalling that $H(x,p) = p \cdot q(x,p) - L(x, q(x,p))$, we see that H is at least of class $C^1(\Omega \times \mathbb{R}^n)$. Consider now $x \in \Omega$, $y \in \mathbb{R}^n$ and $\lambda > 0$ small enough. We have

$$\begin{aligned}
&\frac{H(x+\lambda y, p) - H(x,p)}{\lambda} \\
&\geq \frac{p \cdot q(x,p) - L(x+\lambda y, q(x,p)) - p \cdot q(x,p) + L(x, q(x,p))}{\lambda} \\
&= \frac{L(x, q(x,p)) - L(x+\lambda y, q(x,p))}{\lambda}.
\end{aligned}$$

Letting $\lambda \to 0$ we obtain

$$\frac{\partial H}{\partial x}(x,p) \cdot y \geq -\frac{\partial L}{\partial x}(x, q(x,p)) \cdot y, \qquad \forall y \in \mathbb{R}^n,$$

which implies (3.7). The C^2 regularity of H is then a consequence of (3.6), (3.7). ∎

Bibliography

[1] Ambrosetti A., Prodi G., *A primer of nonlinear analysis*, Cambridge University Press, Cambridge (1993).

[2] Bellman R., *Dynamic programming*, Princeton University Press (1957).

[3] Buttazzo G., Giaquinta M., Hildebrandt S., *One-dimensional Variational Problems*, Clarendon Press, Oxford (1998).

[4] Carathéodory C., *Calculus of variations and partial differential equations of the first order*, Teubner, Berlin (1935).

[5] Chiang A. C., *Elements of Dynamic Optimization*, Mc Graw–Hill International Editions, Economic Series (1992).

[6] Clarke F. H.,*Optimization and nonsmooth analysis*, Wiley, New York (1983).

[7] Clarke F. H., Ledyaev Y. S., Stern R. J., Wolenski P.R., *Nonsmooth analysis and control theory*, Graduate Texts in Mathematics, Springer, New York (1998).

[8] Conway E. D., Hopf E., Hamilton's theory and generalized solutions of the Hamilton-Jacobi equation, *J. Math. Mech.* **13** (1964), 939–986.

[9] Douglis A., The continuous dependence of generalized solutions of non–linear partial differential equations upon initial data, *Comm. Pure Appl. Math.* **14** (1961), 267–284.

[10] Eisner R., Strotz R. H., *Determinants of business investment*, Impacts of Monetary Policy, A Series of Research Studies Prepared for the Commition on Money and Credit, Prentice–Hall, Englewood Cliffs, NJ, 1963, 60–233.

[11] Evans G. C., *The dynamics of monopoly*, American Mathematical Monthly, February 1924, 77–83.

[12] Evans L. C., *Partial Differential Equations*, A.M.S. (1998).

[13] Evans L.C., Gariepy R. F., *Measure theory and fine properties of functions*, Studies in Advanced Mathematics, CRC Press, Ann Arbor (1992).

[14] Demange G., Rochet J. C. , *Methodes Mathematiques de la Finance*, Economica (1997).

[15] Fattorini H. O., *Infinite Dimensional Optimization and Control Theory*, Encyclopedia of Mathematics and Its Applications, Cambridge University Press (1998).

[16] Hamermesh D. S., *Labor demand and the structure of adjustment costs*, American Economic Review, September 1989, 674–689.

[17] Hotelling H., *The Economics of Exhaustible Resources*, Journal of Political Economy, April 1931, 137–175.

[18] Nordhaus W. D., *The political business cycle*, Review of Economic Studies, April 1975, 169–190.

[19] Rockafellar R. T., *Convex Analysis*, Princeton University Press, Princeton (1970).

[20] Rockafellar R. T., Wets R. J.-B.,*Variational Analysis*, Grundlehren der Math. Wiss., Springer-Verlag, Berlin (1998).

[21] Taylor D., *Stopping inflation in the Dornbusch model: optimal monetary policies with alternate price–adjustment equations*, Journal of Macroeconomics, Spring 1989, 199–216.

[22] Troutman J. L. *Variational Calculus and Optimal Control*, Undergraduate Texts in Mathematics, Springer, New York (1998).

[23] Zajíček L., On the differentiation of convex functions in finite and infinite dimensional spaces, *Czechoslovak Math. J.* **29** (1979), 340–348.