OPTIMAL CONTROL SYSTEMS WITH TIME DELAY AND CONDITIONS FOR NORMALITY

by

Javier Fernando Rosenblueth Laguette

A thesis submitted for the Degree of
Doctor of Philosophy of the University of London
and for the
Diploma of Membership of
the Imperial College of Science and Technology

November 1986

DEPARTMENT OF ELECTRICAL ENGINEERING

IMPERIAL COLLEGE OF SCIENCE AND TECHNOLOGY

UNIVERSITY OF LONDON

ABSTRACT

We treat two fundamental problems in dynamic optimization. The first one deals with the fixed endpoint problem in the calculus of variations involving a delay in the phase coordinates. Necessary conditions in the form of a maximum principle are well known and, hence, conditions equivalent to those of Euler, Legendre and Weierstrass. However, no results seem to exist for sufficiency or for a corresponding Jacobi condition. We derive necessary and sufficient conditions in terms of the first and second variations, extending the classical results for the delay free case. The first order condition is then characterized in terms of Euler's equation and, for the second order condition, we obtain through the method of steps and solutions of the Hamiltonian system conditions similar to that of Jacobi. The second problem, normality for optimal control systems (the cost or performance index in the necessary conditions of Pontryagin's maximum principle does not vanish) is studied through perturbations of the endpoint set. It is known that normality holds for all problems obtained by translating the original endpoint set in directions belonging to a dense set. When the equations of motion are linear in the state variable, we enlarge this set of directions to a full (Lebesgue) measure set. For non-linear systems, we show that enlarging or diminishing the endpoint set, instead of translating it, normality is also guaranteed almost everywhere.

CONTENTS

ACKNOWLEDGEMENTS

I am most grateful to Dr. R. B. Vinter. This research was done under his supervision and I wish to express my many thanks for all his advice and help.

I am greatly indebted to the Consejo Nacional de Ciencia y Tecnologia, Mexico, for the financial support.

And to everyone: the encouragement has been permanent.

"שושנה: סתירה מושלמת.
הנאה לא להיות החלום של אף אחד
בין כל כך הרבה עיניים עצומות".

To

Veronika E. Rohen

PART I

OPTIMAL CONTROL SYSTEMS

WITH TIME DELAY

CHAPTER 1

THE DELAY FREE PROBLEM

1 INTRODUCTION

In this chapter we present a brief summary of the classical theory on necessary and sufficient conditions for the simple fixed endpoint problem in the calculus of variations. These conditions are first given in a non-standard form involving the first and second variations. The standard form, in terms of the Euler equation and the Jacobi condition, is then derived by characterizing the previous first and second-order conditions and, independently, making use of the field theory and the Hamilton-Jacobi inequality.

The purpose of presenting three different proofs of the same theorem is twofold. First of all, we want to see how much of the classical theory can be extended in a natural way to the problem involving delays. It will turn out to be the non-standard form of the conditions the only approach where we can put in parallel both theories, word by word. Difficulties arise in trying to characterize the second-order condition and finding an equivalence for Jacobi's condition. We shall see all this in detail in Chapter 2.

On the other hand, we convert the delay problem, in Chapter 3, into a non-delay problem. It will no longer have fixed endpoint constraints, but several results of the classical theory will serve as the main tool for finding the corresponding necessary and sufficient conditions.

Most results in this chapter are based on the classical literature (see, for example, [4], [5], [15], [18], [20], [22], [33]). Some proofs are simplified and we make weaker assumptions than usual. For some reason, there are no common assumptions concerning the smoothness of the functions delimiting the problem. To give an example, in order to apply the classical sufficient conditions, the Lagrangian is assumed to be of class C^4 in [15] and [20] and of class C^2 in [11] and [18]. In [15] and [20], the trajectory under consideration is assumed to be of class C^4, in [18] of class C^1 and in [11] piecewise-C^1. The assumptions seem to be, for each case, an integral component of the proofs.

What we show is that, for each specific result, it is enough to impose the continuity of the Lagrangian and the partial derivatives involved. In particular, the sufficient conditions hold if the Lagrangian $L(t,x,\dot{x})$ and its first and second partial derivatives with respect to x and $\dot{x}$ are continuous.

For this theorem, we also prove that it suffices for the trajectory to be piecewise-smooth as in [11], but the conditions (either for a weak or a strong minimum) do not only imply that it solves the problem locally: if the trajectory satisfies these conditions, it has to be continuously differentiable.

In Chapter 2, the Lagrangian will be given by a function $L(t,x,u,\dot{x})$ mapping $[t_0,t_1] \times R^{3n}$ to R. If L does not depend on the variable u and the delay is zero, the problem will be reduced to the non-delay case. Consequently, those propositions and theorems whose proofs are essentially the same for both problems, will be stated in this chapter without proof, and most comments and observations will be left for the delay problem.

2 STATEMENT OF THE PROBLEM AND VARIATIONS AS DIFFERENTIALS

We are given an interval $[t_0,t_1]$, two points ξ_0 and ξ_1 in R^n, an open set A in Ψ, where Ψ denotes the usual topology of $[t_0,t_1] \times R^{2n}$, and a function L mapping $[t_0,t_1] \times R^{2n}$ to R (called the **Lagrangian**). Let

$$X := \{x:[t_0,t_1] \to R^n \mid x \text{ is piecewise-}C^1\}$$

(space of trajectories),

$$X(A) := \{x \in X \mid \tilde{x}(t) \in A,\ (t_0 \leqslant t \leqslant t_1),\ \text{and } L \circ \tilde{x} \text{ is integrable}\}$$

(admissible trajectories)

$$X_{\bullet}(A) := \{x \in X(A) \mid x(t_0) = \xi_0 \text{ and } x(t_1) = \xi_1\}$$

(endpoint constraints)

where, for all x in X and t in $[t_0,t_1]$, $\tilde{x}(t) := (t,x(t),\dot{x}(t))$. The basic problem in the calculus of variations, which we label P(A), is that of minimizing the functional $I(x) := \int_{t_0}^{t_1} L(\tilde{x}(t))dt$ over $X_{\bullet}(A)$. Of course, the problem depends on the Lagrangian, the endpoint constraints and the set A but, for most purposes it will be convenient to write explicitly only its dependence on A.

A "solution" of P(A) is considered in the following senses: for all x in X and $\varepsilon > 0$, let

$$T_0(x;\varepsilon) := \{(t,y) \in [t_0,t_1] \times R^n \mid |x(t) - y| < \varepsilon\}$$

(tube about x)

$$T_1(x;\varepsilon) := \{(t,y,v) \in T_0(x;\varepsilon) \times R^n \mid |\dot{x}(t) - v| < \varepsilon\}$$

(restricted tube about x).

A trajectory x is said to solve P(A) if

$$x \in S(A) := \{x \in X_{\bullet}(A) \mid I(x) \leqslant I(y) \quad \forall\, y \in X_{\bullet}(A)\},$$

x is a **strong minimum for P(A)** if there exists $\varepsilon > 0$ such that x solves $P((T_0(x;\varepsilon) \times R^n) \cap A)$ and x is a **weak minimum for P(A)** if there exists $\varepsilon > 0$ such that x solves $P(T_1(x;\varepsilon) \cap A)$.

We introduce next some definitions in terms of the space X and the Lagrangian that will be used constantly. Let us denote the values of L by $L(t,x,\dot{x})$. In (ii)-(iv), the domain of the functions is defined wherever the derivatives involved make sense.

i. **Set of trajectories vanishing at t_0 and t_1.**

$$Y := \{y \in X \mid y(t_0) = y(t_1) = 0\}.$$

ii. **The Weierstrass 'excess function'.**

$$E(t,x,\dot{x},u) := L(t,x,u) - L(t,x,\dot{x}) - \langle u - \dot{x} \,,\, L_{\dot{x}}(t,x,\dot{x})\rangle.$$

iii. **The first variation of L with respect to x ε X.**

$$I'(x;y) := \int_{t_0}^{t_1} (\langle L_x(\tilde{x}(t)) \,,\, y(t)\rangle + \langle L_{\dot{x}}(\tilde{x}(t)) \,,\, \dot{y}(t)\rangle)dt \quad \text{for all } y \in X.$$

iv. **The second variation of L with respect to x ε X.**

$$I''(x;y) := \int_{t_0}^{t_1} 2\Omega(t,y(t),\dot{y}(t))dt \qquad \text{for all } y \in X$$

where, for all $(t,y,\dot{y})$ in $[t_0,t_1] \times R^{2n}$,

$$2\Omega(t,y,\dot{y}) := \langle y \,,\, L_{xx}(\tilde{x}(t))y\rangle + 2\langle y \,,\, L_{x\dot{x}}(\tilde{x}(t))\dot{y}\rangle + \langle \dot{y} \,,\, L_{\dot{x}\dot{x}}(\tilde{x}(t))\dot{y}\rangle.$$

We end this section stating an initial result concerning the variations. For ease of notation we shall say from now on that, for example, $L \in C^2(A;x,\dot{x})$, if L and its first and second partial derivatives with respect to x and $\dot{x}$ are continuous on A.

1. Proposition: Assume $L \in C^2(A;x,\dot{x})$. Then the functions $I'(x;\cdot)$ and $I''(x;\cdot)$ are the first and second Fréchet differentials at x of the functional I with respect to the space $(X, \|\ \|)$ where, for all x in X, $\|x\| := \sup\{|x(t)| + |\dot{x}(t)| \mid t_0 \leqslant t \leqslant t_1\}$.

An immediate consequence of this fact are the following sufficient conditions for a weak minimum.

2. Corollary: Suppose $L \in C^2(A;x,\dot{x})$ and x_0 is a trajectory satisfying the endpoint constraints. If there exists $\varepsilon > 0$ such that $I'(x_0;y) = 0$ and $I''(x;y) > 0$ for all y in $Y - \{0\}$ and all x in $X_\bullet(A)$ satisfying $\|x - x_0\| < \varepsilon$, then x_0 is a strict weak minimum for P(A).

3 NECESSARY AND SUFFICIENT CONDITIONS THROUGH VARIATIONS

The purpose of this section is to state the classical necessary and sufficient conditions for a minimum in terms, explicitly, of the first and second variations. For all A in $\mathcal{Y}$, consider the following sets:

(1) **Vanishing of the first variation in Y.**

$$E(A) := \{x \in X(A) \mid I'(x;y) = 0 \quad \forall\, y \in Y\}$$

(called **the set of extremals**).

(2) **Nonnegativity of the second variation in Y.**

$$H(A) := \{x \in X(A) \mid I''(x;y) \geqslant 0 \quad \forall\, y \in Y\}.$$

(3) **The condition of Weierstrass.**

$$W(A) := \{x \in X(A) \mid E(t,x(t),\dot{x}(t),u) \geqslant 0 \quad \forall\ (t,x(t),u) \in A\}.$$

(4) **The condition of Legendre.**

$$L(A) := \{x \in X(A) \mid L_{\dot{x}\dot{x}}(\tilde{x}(t)) \geq 0 \quad \forall\, t \in [t_0, t_1]\}.$$

With these definitions, the following result holds:

3. **Theorem:** Suppose L is continuous on A and x is a trajectory solving P(A). Then x lies in E(A), H(A), W(A) and L(A) whenever the derivatives of L involved (in each case) are continuous. In particular, if $L \in C^2(A;x,\dot{x})$, then:

$$S(A) \subset E(A) \cap H(A) \cap W(A) \cap L(A).$$

Consider now the following sets of trajectories, obtained by slightly strengthening the previous ones.

(2)′ **Positivity of the second variation in Y − { 0 }.**

$$H'(A) := \{x \in H(A) \mid I''(x;y) > 0 \quad \forall\, y \in Y - \{0\}\}.$$

(3)′ **Strengthened Weierstrass condition.**

$$W(A;\varepsilon) := \{x_0 \in W(A) \mid E(t,x,\dot{x},u) \geq 0 \quad \forall\, (t,x,\dot{x}) \in T_1(x_0;\varepsilon) \text{ and } (t,x,u) \in A\}.$$

(4)′ **Strengthened Legendre condition.**

$$L'(A) := \{x \in L(A) \mid L_{\dot{x}\dot{x}}(\tilde{x}(t)) > 0 \quad \forall\, t \in [t_0, t_1]\}.$$

This slight strengthening of the necessary conditions is adequate to render them sufficient, as the following theorem shows.

4. **Theorem:** Suppose $L \in C^2(A;x,\dot{x})$ and x_0 is a trajectory satisfying the endpoint constraints.

i. If x_0 lies in E(A), H′(A) and L′(A), then x_0 is a (strict) weak minimum for P(A) of class $C^1([t_0,t_1])$.

ii. If, for some $\varepsilon > 0$, x_0 lies in E(A), H′(A), W(A;ε) and L′(A), then x_0 is a (strict) strong minimum for P(A) of class $C^1([t_0,t_1])$.

Several results hold if, instead of L′(A), we impose along a trajectory only the nonsingularity of $L_{\dot{x}\dot{x}}$. Consider the following strengthened set of extremals:

(1)′ **Nonsingular extremals.**

$$E'(A) := \{x \in E(A) \mid |L_{\dot{x}\dot{x}}(\tilde{x}(t))| \neq 0 \quad \forall\ t \in [t_0,t_1]\}$$

and observe that Theorem 4 can be restated as follows:

i. Suppose $x \in X_\bullet(A) \cap E'(A) \cap H'(A) \cap L(A)$. Then there exists $\varepsilon > 0$ such that $x \in S(T_1(x;\varepsilon) \cap A) \cap C^1([t_0,t_1])$.

ii. Suppose $x \in X_\bullet(A) \cap E'(A) \cap H'(A) \cap W(A;\varepsilon)$, for some $\varepsilon > 0$. Then there exists $\delta > 0$ such that $x \in S((T_0(x;\delta) \times R^n) \cap A) \cap C^1([t_0,t_1])$.

It is easy to show that, if (ii) is true, then the assertion in (i) holds. This fact, which will be proved in Chapter 2, allows us to study these sufficient conditions only for a strong minimum.

4 THE FIRST VARIATION: EULER'S EQUATION

The whole purpose of the remaining sections is to establish the classical necessary and sufficient conditions in the standard form, i.e., in terms exclusively of the Lagrangian without referring to the variations. We start characterizing the set of extremals, that is, the set E(A) of trajectories where the first variation vanishes on Y, and prove that it is equivalent to the set of trajectories satisfying the so-called Euler's equation. This fact, together with some properties of extremals, will follow again as a particular case of the problem with delays.

5. **Proposition:** Suppose x is a trajectory in X(A). Then x is an extremal if, and only if, there exists a constant c in R^n such that, for all t in $[t_0,t_1]$,

$$L_{\dot{x}}(\tilde{x}(t)) = \int_{t_0}^{t} L_x(\tilde{x}(s))ds + c.$$

This last relation is known as **Euler's equation.** It has several consequences and we summarize the most important ones in the following two corollaries. It should be emphasized that the smoothness assumptions are weaker than usual and that the second corollary, which is not a standard result, will permit us to assume proving the classical sufficient conditions that the trajectory under consideration is only piecewise-C^1.

6. **Corollary:** For any set A in Ψ, the following holds:

i. **Weierstrass-Erdmann corner condition.** If x ε E(A) then $L_{\dot{x}} \circ \tilde{x}$ is continuous on $[t_0,t_1]$.

ii. **Hilbert Differentiability Theorem.** If L_x is $C^{r-2}(A)$ and $L_{\dot{x}}$ is $C^{r-1}(A)$ $(r \geq 2)$, then any nonsingular extremal of class C^1 is of class C^r, i.e., $E'(A) \cap C^1([t_0,t_1]) \subset C^r([t_0,t_1])$.

iii. **Regularity.** If L is **regular on A** (that is, if A is convex in $\dot{x}$ and $L_{\dot{x}\dot{x}}(t,x,\dot{x}) > 0$ for all $(t,x,\dot{x})$ in A) and $L_{\dot{x}\dot{x}}$ is continuous on A, then:

a. $E(t,x,\dot{x},u) > 0$ for all $(t,x,\dot{x}) \in A$ and $(t,x,u) \in A$.

b. $E(A) \subset C^1([t_0,t_1])$.

7. **Corollary:** Suppose L and $L_{\dot{x}\dot{x}}$ are continuous on A, x_0 is a nonsingular trajectory (that is, $|L_{\dot{x}\dot{x}}(\tilde{x}(t))| \neq 0$ for all t in $[t_0,t_1]$) and, for some ε > 0, x_0 belongs to W(A;ε). Then ε can be diminished so

that the inequality in the definition of $W(A;\varepsilon)$ becomes strict, i.e., $E(t,x,\dot{x},u) > 0$ for all $(t,x,\dot{x})$ in $T_1(x_0;\varepsilon)$ and (t,x,u) in A with $u \neq \dot{x}$. Moreover, if x_0 is also an extremal, then $x_0 \in C^1([t_0,t_1])$. In other words, for all $\varepsilon > 0$, $E'(A) \cap W(A;\varepsilon) \subset C^1([t_0,t_1])$ and, in particular, $E'(A) \cap L(A) \subset C^1([t_0,t_1])$.

As we shall see in a moment, nonsingular extremals satisfy a system of ordinary differential equations. This is the central result in the calculus of variations and it should be noted that practically all results we state from now on, rely on this fact (which is a simple consequence of the implicit function theorem) and the classical theory on existence and uniqueness of solutions of this kind of systems. It is also the main difference between this and the delay problem since, for the latter one, nonsingular extremals satisfy instead a system of differential difference equations involving both advanced and retarded arguments.

8. Proposition: Suppose $L_{\dot{x}}$ is of class $C^r(A)$ $(r \geq 1)$, and x_0 is a nonsingular trajectory in $X(A)$. Set $p_0(t) := L_{\dot{x}}(\tilde{x}_0(t))$ and, for all $\varepsilon > 0$, let $T_2(x_0;\varepsilon) := \{(t,x,p) \in T_0(x_0;\varepsilon) \times \mathbf{R}^n \mid |p - p_0(t)| < \varepsilon\}$. Then there exist $\varepsilon > 0$ and a function $\Lambda(t,x,p)$, with Λ_p of class C^r on $T_2(x_0;\varepsilon)$, such that the following are equivalent:

i. $x \in E'(A) \cap C^1([t_0,t_1])$, $p(t) = L_{\dot{x}}(\tilde{x}(t))$ and $(t,x(t),p(t))$ belongs to $T_2(x_0;\varepsilon)$.

ii. (x,p) satisfies the ordinary differential system:

$$\text{(EE)} \qquad \begin{cases} \dot{x}(t) = \Lambda_p(t,x(t),p(t)) \\ \dot{p}(t) = -\Lambda_x(t,x(t),p(t)) \end{cases}$$

Proof: Let $G(t,x,p,v) := L_{\dot{x}}(t,x,v) - p$ for all t in $[t_0,t_1]$ and x, p and v in $\mathbf{R}^n$. By definition of p_0, $G(t,x_0(t),p_0(t),\dot{x}_0(t)) = 0$ and, since x_0 is nonsingular, $|G_v(t,x_0(t),p_0(t),\dot{x}_0(t))| \neq 0$. $L_{\dot{x}}$ is C^r and so is G.

Consequently, by the implicit function theorem, there exist $\varepsilon > 0$ and a unique function U of class C^r mapping $T_2(x_0;\varepsilon)$ to R^n, such that, for all (t,x,p) in $T_2(x_0;\varepsilon)$, $p = L_{\dot{x}}(t,x,U(t,x,p))$. Define for all (t,x,p) in $T_2(x_0;\varepsilon)$:

$$\Lambda(t,x,p) := \langle p\ ,\ U(t,x,p)\rangle - L(t,x,U(t,x,p)).$$

This function is called the **Legendre transform** (with respect to x_0) and observe that it satisfies:

$$\Lambda_x(t,x,p) = -\ L_x(t,x,U(t,x,p))$$

$$\Lambda_p(t,x,p) = U(t,x,p).$$

This implies the required smoothness of Λ. Now, clearly, (i) $\Rightarrow$ (ii) follows by uniqueness of U and Proposition 5 and (ii) $\Rightarrow$ (i) follows diminishing ε so that $|L_{\dot{x}\dot{x}}(t,x,\dot{x})| \neq 0$ for all $(t,x,\dot{x})$ in $T_1(x_0;\varepsilon)$.

Remark: Throughout the remaining sections, we shall need the assumptions $L \in C^2(A;x)$ and $L_{\dot{x}} \in C^1(A)$. The proof of Proposition 8 shows not only that, for this case, Λ_p is C^1, but also that Λ_x inherits the smoothness of L_x, i.e., Λ is also $C^2(T_2(x_0;\varepsilon);x)$.

5 THE SECOND VARIATION: JACOBI'S CONDITION

In this section we characterize, as we did with the set of extremals, the set of trajectories where the second variation is nonnegative in Y (H(A)) or strictly positive (H'(A)). We start with a few definitions.

Given a trajectory x, denote by J_x the second variation with respect to x and by E_x the set of trajectories (called **secondary extremals**), that satisfy Euler's equation for the integrand Ω, i.e., for all y in X,

$$J_x(y) := I''(x;y) = \int_{t_0}^{t_1} 2\Omega(t, y(t), \dot{y}(t))dt$$

$$E_x := \{y \in X \mid J_x'(y;z) = 0 \quad \forall\ z \in Y\}$$

$$= \{y \in X \mid \text{there exists } c \in \mathbf{R}^n \text{ such that}$$

$$\Omega_{\dot{y}}(\tilde{y}(t)) = \int_{t_0}^{t} \Omega_y(\tilde{y}(s))ds + c \quad \forall\ t_0 \leqslant t \leqslant t_1\}.$$

Secondary extremals can be characterized in a simple way in terms of a linear system in $\mathbf{R}^n$. It corresponds, for the integrand Ω, to the system we already met in Proposition 8. Observe first that, given x in X,

$$\Omega_y(t, y, \dot{y}) = L_{xx}(\tilde{x}(t))y + L_{x\dot{x}}(\tilde{x}(t))\dot{y}$$

$$\Omega_{\dot{y}}(t, y, \dot{y}) = L_{\dot{x}x}(\tilde{x}(t))y + L_{\dot{x}\dot{x}}(\tilde{x}(t))\dot{y}.$$

Now, if $x \in L'(A)$, the integrand Ω is regular and, by Corollary 6(iii), all secondary extremals (with respect to x) are of class C^1. This implies the following result:

9. Proposition: Suppose $x \in L'(A)$ and $y \in X$. Then the following are equivalent:

i. y is a secondary extremal (with respect to x).

ii. For all t in $[t_0, t_1]$, y satisfies the differential equation:

$$\text{(JE)} \qquad \frac{d}{dt}[L_{\dot{x}x}(\tilde{x}(t))y(t) + L_{\dot{x}\dot{x}}(\tilde{x}(t))\dot{y}(t)] = L_{xx}(\tilde{x}(t))y(t) + L_{x\dot{x}}(\tilde{x}(t))\dot{y}(t).$$

called the **Jacobi equation.**

iii. If we set $q(t) := L_{\dot{x}x}(\tilde{x}(t))y(t) + L_{\dot{x}\dot{x}}(\tilde{x}(t))\dot{y}(t)$, then (y,q) satisfies the linear system (* denotes transpose):

$$\text{(JE)}' \qquad \begin{cases} \dot{y}(t) = A(t)y(t) + B(t)q(t) \\ \dot{q}(t) = C(t)y(t) - A^*(t)q(t) \end{cases}$$

where:

$$A(t) := -L_{\dot{x}\dot{x}}^{-1}(\tilde{x}(t))L_{\dot{x}x}(\tilde{x}(t)),$$

$$B(t) := L_{\dot{x}\dot{x}}^{-1}(\tilde{x}(t))$$

$$C(t) := L_{xx}(\tilde{x}(t)) - L_{x\dot{x}}(\tilde{x}(t))L_{\dot{x}\dot{x}}^{-1}(\tilde{x}(t))L_{\dot{x}x}(\tilde{x}(t)).$$

Now, assuming $x \in L'(A) \cap C^1([t_0,t_1])$ and $L \in C^2(A;x,\dot{x})$, the matrices A, B and C are continuous and the theorem on the existence and uniqueness of the solution of the Cauchy problem holds for (JE)′. Denote by $(Y(\cdot,t_0),Q(\cdot,t_0))$ the fundamental (matrix) solution of the Jacobi equation, i.e., the (matrix) solution of (JE)′ satisfying the initial data:

$$Y(t_0,t_0) = 0, \qquad Q(t_0,t_0) = I.$$

A point s in $(t_0,t_1]$ is called **conjugate to t_0 with respect to x** if $Y(s,t_0)$ is degenerate. Observe that this is equivalent to the existence of a secondary extremal y in E_x nonvanishing on (t_0,s) and such that $y(t_0) = y(s) = 0$.

This definition allows us to express C^1 trajectories in H(A), satisfying the strengthened Legendre condition, in terms of the Jacobi equation. Consider the following set:

(5) Jacobi's condition.

$$J(A) := \{x \in X(A) \mid s \in (t_0,t_1) \Rightarrow s \text{ is not conjugate to } t_0 \text{ with respect to } x\}.$$

10. Proposition: Suppose $L \in C^2(A;x,\dot{x})$ and x is a C^1 trajectory in H(A) satisfying the strengthened Legendre condition. Then there are no conjugate points to t_0 with respect to x in the open interval (t_0,t_1). In terms of (5), $C^1([t_0,t_1]) \cap L'(A) \cap H(A) \subset J(A)$.

Proof: Suppose $x \notin J(A)$. By definition, there exist s in (t_0,t_1) and y in E_x, such that y is not identically zero on (t_0,s) and $y(t_0) = y(s) = 0$. Let $z(t) := y(t)$ for all t in $[t_0,s]$ and $z(t) := 0$ for all t in $[s,t_1]$. Clearly z belongs to Y and

$$J_x(z) = \int_{t_0}^{s} 2\Omega(t,z(t),\dot{z}(t))dt$$

$$= \int_{t_0}^{s} \{\langle z(t) \ , \ \Omega_y(\tilde{z}(t))\rangle + \langle \dot{z}(t) \ , \ \Omega_{\dot{y}}(\tilde{z}(t))\rangle\}dt$$

$$= \int_{t_0}^{s} \frac{d}{dt}\{\langle z(t) \ , \ \Omega_{\dot{y}}(\tilde{z}(t))\rangle\}dt = 0.$$

Since $x \in H(A)$, z minimizes J_x on Y. Now, since x is C^1, Ω, Ω_y, and $\Omega_{\dot{y}}$ are continuous and so, by Theorem 3, $z \in E_x$. But this implies, since $z(t) = \dot{z}(t) = 0$ for all t in (s,t_1) and z satisfies (JE), that $z(t) \equiv 0$ on $[t_0,t_1]$. This contradicts the nonvanishing of y and the result follows.

This proposition and Theorem 3 imply the classical necessary conditions for a minimum in the standard form. Observe that we are not assuming x is C^1. This follows by the conditions and Corollary 7.

11. **Theorem:** Suppose $L \in C^2(A;x,\dot{x})$ and x is a trajectory solving P(A). Then x satisfies Euler's equation and the conditions of Legendre and Weierstrass. If also x is nonsingular then x satisfies Jacobi's condition.

We have thus expressed the set H(A) in terms of the Jacobi equation.

We turn now to the set $H'(A)$. As we shall see below, it can be characterized in terms of conjugate points imposing the same condition as for $H(A)$ but in the half-open interval $(t_0,t_1]$. It should be noted that, for this characterization, the smoothness assumptions concerning the Lagrangian will be slightly stronger than $L \in C^2(A;x,\dot{x})$: we also require for $L_{\dot{x}}$ to be C^1 on A.

(5)′ Strengthened Jacobi's condition.

$$J'(A) := \{x \in J(A) \mid t_1 \text{ is not conjugate to } t_0 \text{ with respect to } x\}.$$

12. **Proposition:** Suppose $L \in C^2(A;x)$, $L_{\dot{x}} \in C^1(A)$ and x_0 is a nonsingular extremal satisfying the Legendre condition. Then the following are equivalent:

i. $x_0 \in H'(A)$.

ii. $x_0 \in J'(A)$.

iii. There exists a matrix solution (Y,Q) of the linear system (JE)′, satisfying $|Y(t)| \neq 0$ and $Y^*(t)Q(t) = Q^*(t)Y(t)$ for all t in $[t_0,t_1]$.

Proof: (i) ⇒ (ii): The assumptions imply, by Corollary 7, that x_0 is C^1 and, applying Proposition 10, $x_0 \in J(A)$. So, if we suppose that (ii) is false, there will exist y in Y, a nonvanishing secondary extremal with respect to x_0. As before, we obtain $I''(x_0;y) = 0$, thus contradicting (i).

(ii) ⇒ (iii): Let $p_0(t) = L_{\dot{x}}(\tilde{x}_0(t))$. By Proposition 8, (x_0,p_0) satisfies (EE), and so it can be extended over a larger interval $[t_0-\varepsilon,t_1+\varepsilon]$. Solve the Cauchy problem for (EE) with initial data $(0 < \delta < \varepsilon)$:

$$x(t_0-\delta,\lambda) = x_0(t_0-\delta)$$

$$p(t_0-\delta,\lambda) = \lambda + p_0(t_0-\delta).$$

By their definition, $x(\cdot,\lambda)$ are extremals of L and, therefore, the following identity holds:

$$-\frac{d}{dt}[L_{\dot{x}}(t,x(t,\lambda),\dot{x}(t,\lambda))] + L_x(t,x(t,\lambda),\dot{x}(t,\lambda)) \equiv 0.$$

Differentiating with respect to λ we obtain, in view of the equality $x(t,0) = x_0(t)$,

$$-\frac{d}{dt}[L_{\dot{x}\dot{x}}(\tilde{x}_0(t))\dot{Y}(t,t_0-\delta) + L_{\dot{x}x}(\tilde{x}_0(t))Y(t,t_0-\delta)]$$

$$+ L_{x\dot{x}}(\tilde{x}_0(t))\dot{Y}(t,t_0-\delta) + L_{xx}(\tilde{x}_0(t))Y(t,t_0-\delta) \equiv 0$$

where $Y(t,t_0-\delta)$ denotes the matrix $\partial x(t,0)/\partial\lambda$. This equation is precisely (JE). So, defining

$$Q(t,t_0-\delta) := L_{\dot{x}\dot{x}}(\tilde{x}_0(t))\dot{Y}(t,t_0-\delta) + L_{\dot{x}x}(\tilde{x}_0(t))Y(t,t_0-\delta),$$

it follows that $(Y(t,t_0-\delta),Q(t,t_0-\delta))$ solves (JE)′.

Observe now that $Q(t,t_0-\delta) = \partial p(t,0)/\partial\lambda$ since, by Proposition 8, $p(t,\lambda) = L_{\dot{x}}(\tilde{x}(t,\lambda))$. So, by the boundary conditions imposed for $x(\cdot,\lambda)$ and $p(\cdot,\lambda)$,

$$Y(t_0-\delta,t_0-\delta) = 0, \quad Q(t_0-\delta,t_0-\delta) = I.$$

This implies that $(Y(t,t_0-\delta),Q(t,t_0-\delta))$ is the fundamental matrix solution of the Jacobi equation. Since we are assuming $x_0 \in J'(A)$, it follows from the continuous dependence of solutions on initial data, that the matrix $Y(t,t_0-\delta)$ is nonsingular for some $\delta > 0$ on the entire interval $[t_0,t_1]$. This completes the first part of (iii). The second part follows from the fact that the quantity $Y^*(t)Q(t) - Q^*(t)Y(t)$ has zero derivative and value 0 at $t = t_0-\delta$.

(iii) $\Rightarrow$ (i): Let $y \in Y$ and set $w(t) := Y^{-1}(t)y(t)$ and $z(t) := Y(t)\dot{w}(t)$ so that $\dot{y}(t) = \dot{Y}(t)w(t) + z(t)$. We want to show that $I''(x_0;y) > 0$ unless $y \equiv 0$. From the definition of Ω,

$$\begin{aligned} 2\Omega(\tilde{y}(t)) &= \langle y(t) , \Omega_y(\tilde{y}(t))\rangle + \langle \dot{y}(t) , \Omega_{\dot{y}}(\tilde{y}(t))\rangle \\ &= \langle Y(t)w(t) , \dot{Q}(t)w(t) + L_{x\dot{x}}(\tilde{x}_0(t))z(t)\rangle \\ &\qquad + \langle \dot{Y}(t)w(t) + z(t) , Q(t)w(t) + L_{\dot{x}\dot{x}}(\tilde{x}_0(t))z(t)\rangle \\ &= \frac{d}{dt}(\langle w(t) , Q^*(t)Y(t)w(t)\rangle) + \langle z(t) , L_{\dot{x}\dot{x}}(\tilde{x}_0(t))z(t)\rangle. \end{aligned}$$

Since $w(t_0) = w(t_1) = 0$, this implies that

$$I''(x_0;y) = \int_{t_0}^{t_1} \langle z(t) , L_{\dot{x}\dot{x}}(\tilde{x}_0(t))z(t)\rangle dt.$$

Consequently, $I''(x_0;y) > 0$ unless $z(t) = Y(t)\dot{w}(t) = 0$. But $|Y(t)| \neq 0$ and $w(t_0) = 0$, so $I''(x_0;y) = 0 \Rightarrow w \equiv 0 \Rightarrow y \equiv 0$. This completes the proof.

The classical sufficient conditions for a minimum follow from this proposition and Theorem 4:

13. **Theorem:** Suppose $L \in C^2(A;x)$, $L_{\dot{x}} \in C^1(A)$ and x in $X_\bullet(A)$ is a non-singular trajectory satisfying Euler's equation, Legendre's condition and Jacobi's strengthened condition. Then x is a weak minimum for P(A). If also L is regular on A or x satisfies the strengthened condition of Weierstrass, then x is a strong minimum for P(A).

6 FIELDS OF EXTREMALS

In the previous section we showed that, if x is a nonsingular extremal satisfying Legendre's condition, then $x \in H'(A)$ if, and only if, $x \in J'(A)$. This allowed us to replace condition (2)' by (5)' in Theorem 4 and we obtained Theorem 13, the standard form of the classical sufficient conditions.

In this and the next section, we prove Theorem 13 directly, that is, without the help of condition H'(A). In fact, the two approaches we present imply sufficiency in a much simpler way than the variational one. This is because of the nature of the problem, but we will have occasion to see how different things are for the delay problem. The first proof is based on the invariance of a line integral and the family $\{x(\cdot,\lambda)\}$ of extremals we met in Proposition 12.

A couple (Γ,M) is called a **Mayer-field on A** if:

i. M is a region in $R \times R^n$.

ii. $\Gamma:M \to R^n$ is of class C^1 and $(t,x,\Gamma(t,x)) \in A$ for all (t,x) in M.

iii. The line integral (called **Hilbert's integral**):

$$I^*:= \int P(t,x)dx + Q(t,x)dt$$

is independent of the path in M, where:

$$P(t,x):= L_{\dot{x}}(t,x,\Gamma(t,x))$$

and

$$Q(t,x):= L(t,x,\Gamma(t,x)) - \langle\Gamma(t,x) , P(t,x)\rangle.$$

Observe that if x is any trajectory in M and we set $L^*(t,x,\dot{x}):= \langle P(t,x) , \dot{x}\rangle + Q(t,x)$, then

$$I^*(x) = \int_{t_0}^{t_1} L^*(t,x(t),\dot{x}(t))dt$$

is minimized by x over the class of trajectories joining its endpoints. If we assume $L \in C^2(A;x,\dot{x})$, then $L^* \in C^1((M \times R^n);x,\dot{x})$, and we can apply Euler's equation for the integrand L^*, that is, there exists a constant c in R^n such that:

$$c = L_{\dot{x}}^*(\tilde{x}(t)) - \int_{t_0}^{t_1} L_x^*(\tilde{x}(s))ds$$

$$= L_{\dot{x}}(t,x(t),\Gamma(t,x(t)))$$

$$- \int_{t_0}^{t_1} \{L_x(s,x(s),\Gamma(s,x(s))) + P_x(s,x(s))(\dot{x}(s) - \Gamma(s,x(s)))\}ds.$$

This implies that, if x is a solution of $\dot{x}(t) = \Gamma(t,x(t))$ (called an **extremal of the field** (Γ,M)), then x is an extremal of L and $I^*(x) = I(x)$. Observe also that, along any trajectory x in M, we have:

$$I(x) = I^*(x) + \int_{t_0}^{t_1} E(t,x(t),\Gamma(t,x(t)),\dot{x}(t))dt$$

since $E(t,x,\Gamma(t,x),\dot{x}) = L(t,x,\dot{x}) - L^*(t,x,\dot{x})$. These facts give us already a sufficient condition for a strict strong minimum.

14. **Lemma:** Suppose $L \in C^2(A;x,\dot{x})$, (Γ,M) is a Mayer-field, and $E(t,x,\Gamma(t,x),\dot{x}) > 0$ for all (t,x) in M and $(t,x,\dot{x})$ in A with $\dot{x} \neq \Gamma(t,x)$. If x_0 in $X_\bullet(A)$ is an extremal of (Γ,M) then, for any trajectory x in $X_\bullet(A) - \{x_0\}$ lying in M, $I(x) > I(x_0)$.

The main result of this section is the fact that extremals satisfying the strengthened Legendre and Jacobi conditions are extremals of some

Mayer-field:

15. **Lemma:** Suppose $L \in C^2(A;x)$, $L_{\dot{x}} \in C^1(A)$ and x_0 is a trajectory in $E'(A) \cap L(A) \cap J'(A)$. Then there exists a Mayer-field (Γ,M) of which x_0 is an extremal.

Proof: In Proposition 12 we proved the existence of some $a_0 < t_0$ such that $|x_\lambda(t,0)| \neq 0$ for all t in $(a_0,t_1]$, where $(x(\cdot,\lambda),p(\cdot,\lambda))$ satisfy the Cauchy problem for (EE) with initial data:

$$x(a_0,\lambda) = x_0(a_0)$$

$$p(a_0,\lambda) = \lambda + p_0(a_0).$$

By the implicit function theorem, there exist $\varepsilon > 0$ and a unique λ of class C^1 mapping $T_0(x_0;\varepsilon)$ to R^n, such that, for all (t,y) in $T_0(x_0;\varepsilon)$, $y = x(t,\lambda(t,y))$. Set $M := T_0(x_0;\varepsilon)$ and $\Gamma(t,y) := \dot{x}(t,\lambda(t,y))$ for all (t,y) in M. Observe that $\Gamma \in C^1(M)$ and $(t,y,\Gamma(t,y)) \in A$ for all (t,y) in M. By construction, x_0 is an extremal of (Γ,M). It remains only to show that I^* is independent of the path in M.

Let $C \subset M$ be any trajectory of class C^1, parametrized by:

$$C = \{(t(s),w(s)) \mid s_0 \leqslant s \leqslant s_1\}$$

and define, for all $s_0 \leqslant s \leqslant s_1$:

$$u(t,s) := x(t,\lambda(t(s),w(s))) \qquad a_0 \leqslant t \leqslant t(s)$$

and

$$F(s) := \int_{a_0}^{t(s)} L(t,u(t,s),\dot{u}(t,s))dt.$$

Evaluating the integral I^* along C, we get:

$$I^*(C) = \int_{s_0}^{s_1} \{\langle P(t(s),w(s)) , \dot{w}(s)\rangle + Q(t(s),w(s))t'(s)\}ds$$

$$= \int_{s_0}^{s_1} \{\langle L_{\dot{x}}(t(s),u(t(s),s),\dot{u}(t(s),s)) , u_s(t(s),s)\rangle$$

$$+ L(t(s),u(t(s),s),\dot{u}(t(s),s))t'(s)\}ds$$

$$= \int_{s_0}^{s_1} F'(s)ds = F(s_1) - F(s_0)$$

and so I^* is independent of C. This completes the proof.

With the help of these lemmas we obtain Theorem 13 as follows: we are assuming that $L \in C^2(A;x)$, $L_{\dot{x}} \in C^1(A)$ and, for some $\varepsilon > 0$,

$$x_0 \in X_e(A) \cap E'(A) \cap J'(A) \cap W(A;\varepsilon).$$

By Lemma 15, x_0 is an extremal of some Mayer-field (Γ,M). By Corollary 7, ε can be diminished so that the inequality in the definition of $W(A;\varepsilon)$ becomes strict. If necessary, diminish also M so that, for all (t,y) in M, $(t,y,\Gamma(t,y)) \in T_1(x_0;\varepsilon)$. Applying Lemma 14 to (Γ,M) we obtain that, for all x in $X_e((M \times R^n) \cap A) - \{x_0\}$, $I(x) > I(x_0)$. This is the definition of a (strict) strong minimum for P(A).

7 THE HAMILTON-JACOBI THEORY

Sufficiency through the Hamilton-Jacobi theory is based on the existence of a function satisfying some inequality or some partial differential equation. In this section we show how, for both cases, the existence of this "verification" function is implied by the classical sufficient conditions for a minimum.

The first proof is based on the Hamilton-Jacobi differential equation and, though we proceed directly, it is equivalent to the one of the last section. The second one relies on the Hamilton-Jacobi inequality. The verification theorem for this approach is not equivalent to the one of the Hamiltonian, but weaker and, unlike the other proofs, sufficiency is derived without the help of a field of extremals. This last proof is due to F. H. Clarke and V. Zeidan (see [11]).

7.1 THE HAMILTONIAN

Given A in Ψ, the **Young-Fenchel transform** or **Hamiltonian** (relative to A), is given, for all (t,x,p) in $[t_0,t_1] \times R^{2n}$, by:

$$H(t,x,p) := \sup\{\langle p,u\rangle - L(t,x,u) \mid (t,x,u) \in A\}.$$

The verification theorem for problem P(A) states the following:

16. Lemma: Let $x \in X_e(A)$ and suppose there exist $\varepsilon > 0$ and a function $W: T_0(x;\varepsilon) \to R$ of class C^1 such that:

a. For all (t,y) in $T_0(x;\varepsilon)$, $W_t(t,y) + H(t,y,W_y(t,y)) = 0$.

b. For all t in $[t_0,t_1]$,

$$H(t,x(t),W_y(t,x(t))) = \langle W_y(t,x(t)) , \dot{x}(t)\rangle - L(\tilde{x}(t)).$$

Then x is a strong minimum for P(A).

Proof: Let $y \in X_e(A)$ with $(t,y(t)) \in T_0(x;\varepsilon)$ for all t in $[t_0,t_1]$. By (a) and the definition of the Hamiltonian,

$$W_t(t,y(t)) + \langle W_y(t,y(t)) , \dot{y}(t)\rangle - L(\tilde{y}(t)) \leqslant 0$$

and, by (a) and (b),

$$W_t(t,x(t)) + \langle W_y(t,x(t)) , \dot{x}(t)\rangle - L(\tilde{x}(t)) = 0.$$

This implies that $I(y) \geqslant W(t_1,\xi_1) - W(t_0,\xi_0) = I(x)$ and we obtain the required result.

Observe now the following: let us assume that $L \in C^2(A;x)$, $L_{\dot{x}} \in C^1(A)$ and $x_0 \in E'(A) \cap L(A) \cap J'(A)$. Let (Γ,M) be the Mayer-field constructed in Lemma 15, of which x_0 is an extremal, and let us define

$$W(t,y):= \int_{a_0}^{t} L(s,x(s,\lambda(t,y)),\dot{x}(s,\lambda(t,y)))ds.$$

The partial derivatives of W are given, for all (t,y) in M, by:

$$W_t(t,y) = L(t,y,\Gamma(t,y)) - \langle P(t,y) , \Gamma(t,y)\rangle$$

$$W_y(t,y) = L_{\dot{x}}(t,y,\Gamma(t,y)) = P(t,y).$$

The Legendre transform, defined in Section 4, was given by:

$$\Lambda(t,x,p):= \langle p , U(t,x,p)\rangle - L(t,x,U(t,x,p))$$

where U is the unique (local) solution around $(t,x_0(t),p_0(t))$ of $p = L_{\dot{x}}(t,x,U(t,x,p))$. By uniqueness of U, M can be diminished so that $\Gamma(t,y) = U(t,y,P(t,y))$. Consequently, for all (t,y) in M,

a'. $W_t(t,y) + \Lambda(t,y,W_y(t,y)) = 0$

b'. $\Lambda(t,x_0(t),W_y(t,x_0(t))) = \langle W_y(t,x_0(t)) , \dot{x}_0(t)\rangle - L(\tilde{x}_0(t))$

which is precisely Lemma 16 with the Legendre's transform instead of the Hamiltonian. Moreover,

$$E(t,x,U(t,x,p),u) = \Lambda(t,x,p) - \langle p,u\rangle + L(t,x,u).$$

So, for a nonsingular trajectory x_0, both transforms coincide (locally) if, and only if, x_0 satisfies the strengthened Weierstrass condition. This fact, together with Lemma 16, implies Theorem 13.

Remark: It should be noted that in this proof we have implicitly shown the relation between Mayer-fields and the Hamilton-Jacobi partial differential equation: given x_0 in $X(A)$ nonsingular, (Γ,M) is a Mayer-field containing x_0 if, and only if, there exists $W:M \to R$ of class C^1 satisfying (a)′ and $W_y(t,y) = L_{\dot{x}}(t,y,\Gamma(t,y))$ for all (t,y) in M.

7.2 THE HAMILTON-JACOBI INEQUALITY

Let us consider now the following verification theorem:

17. Lemma: Let $x \in X_\bullet(A)$ and suppose there exist $\varepsilon > 0$ and a function $W:T_0(x;\varepsilon) \to R$ of class C^1, such that

$$\text{(HJI)}\qquad \begin{aligned} &W_t(t,y) + \langle W_y(t,y)\ ,\ v\rangle - L(t,y,v) \\ &\qquad \leqslant W_t(t,x(t)) + \langle W_y(t,x(t))\ ,\ \dot{x}(t)\rangle - L(t,x(t),\dot{x}(t)) \end{aligned}$$

for all (t,y,v) in $(T_0(x;\varepsilon) \times R^n) \cap A$. Then x is a strong minimum for $P(A)$.

The assertion is readily verified by simply integrating (HJI). Observe that, if W satisfies the Hamilton-Jacobi differential equation, then W satisfies (HJI) but the converse is not necessarily true.

The third proof is as follows: our assumptions are, once more, that $L \in C^2(A;x)$, $L_{\dot{x}} \in C^1(A)$ and, for some $\eta > 0$,

$$x \in X_\bullet(A) \cap E'(A) \cap J'(A) \cap W(A;\eta).$$

It is not difficult to show (see [11] for details) that, for some $\mu > 0$, there is no nontrivial solution (y,q) of (JE)′ with C replaced by $C - \mu I$, for which y vanishes both at t_0 and at some point s in $(t_0,t_1]$. In view of Proposition 12, there exists a matrix solution (Y,Q) of the system (JE)′ with C replaced by $C - \mu I$, such that $|Y(t)| \neq 0$ and $Y^*(t)Q(t) = Q^*(t)Y(t)$ for all t in $[t_0,t_1]$ (this is is proved in [11] without extending the extremal x.).

Setting $V(t) := Q(t)Y^{-1}(t)$ it follows that V is a symmetric solution on $[t_0, t_1]$ of the matrix Riccati inequality:

$$\dot{V}(t) + V(t)A(t) + A^*(t)V(t) + V(t)B(t)V(t) - C(t) < 0.$$

Define, for all (t,y) in $[t_0, t_1] \times R^n$,

$$W(t,y) := \langle L_{\dot{x}}(\tilde{x}(t)) , y\rangle + \frac{1}{2} \langle y - x(t) , V(t)(y - x(t))\rangle.$$

We shall show that the domain of W can be diminished so that W satisfies (HJI). First of all, observe that, in view of the assumptions, $L_{\dot{x}} \circ \tilde{x}$ and hence W are C^1. By the implicit function theorem, there exist $\delta > 0$ and a unique function $u: T_0(x;\delta) \to R^n$ of class C^1, such that, for all (t,y) in $T_0(x;\delta)$,

$$L_{\dot{x}}(t,y,u(t,y)) \ (= W_y(t,y)) = L_{\dot{x}}(\tilde{x}(t)) + V(t)(y - x(t)).$$

Define next, for all (t,y) in $T_0(x;\delta)$,

$$F(t,y) := W_t(t,y) + \langle W_y(t,y) , u(t,y)\rangle - L(t,y,u(t,y)).$$

Evaluating W_t and W_y, one finds that F is given by:

$$\begin{aligned} F(t,y) = &\langle L_x(\tilde{x}(t)) , y\rangle - \langle \dot{x}(t) , V(t)(y - x(t))\rangle \\ &+ \tfrac{1}{2}\langle y - x(t) , \dot{V}(t)(y - x(t))\rangle - L(t,y,u(t,y)) \\ &+ \langle L_{\dot{x}}(\tilde{x}(t)) + V(t)(y - x(t)) , u(t,y)\rangle. \end{aligned}$$

From this it follows that F_y and F_{yy} are continuous in (t,y) and satisfy

$$F_y(t,x(t)) = 0$$

$$F_{yy}(t,x(t)) = \dot{V}(t) + V(t)A(t) + A^*(t)V(t) + V(t)B(t)V(t) - C(t) < 0.$$

So, by Taylor's formula, δ can be diminished so that, for all (t,y) in $T_0(x;\delta)$, $F(t,y) \leq F(t,x(t))$. Setting $\varepsilon := \min\{\eta, \delta\}$ we obtain the required result for, if (t,y,v) is any point in $(T_0(x;\varepsilon) \times R^n) \cap A$, then

$$W_t(t,y) + \langle W_y(t,y) , v\rangle - L(t,y,v)$$

$$= W_t(t,y) + \langle L_{\dot{x}}(t,y,u(t,y)) , v\rangle - L(t,y,v)$$

$$\leqslant W_t(t,y) + \langle L_{\dot{x}}(t,y,u(t,y)) , u(t,y)\rangle - L(t,y,u(t,y))$$

$$= F(t,y) \leqslant F(t,x(t))$$

$$= W_t(t,x(t)) + \langle W_y(t,x(t)) , \dot{x}(t)\rangle - L(t,x(t),\dot{x}(t)).$$

CHAPTER 2

SYSTEMS WITH ONE DELAY

1 INTRODUCTION

The purpose of this chapter is to generalize the results of Chapter 1 to systems involving one delay in the phase coordinates. In order to visualize the parallel between the latter and delay free systems, we shall use the same notation as in the previous chapter. Every concept is followed implicitly by the words "for systems with one delay".

Sections 2 and 3 are devoted to a detailed study of the first and second variations. They are expressed as differentials and necessary and sufficient conditions are derived extending, word by word, the classical results for the delay free problem. In Section 4 we characterize the set of extremals, finding the corresponding Euler's equation together with several well known consequences of solutions of the equation. Section 5 treats the question of expressing $H'(A)$, the set of trajectories for which the second variation is strictly positive, in terms of solutions of the Jacobi equation. We show how this equation is equivalent to a linear delay system for which, if one imposes conditions both in the initial and terminal intervals of length the delay, existence and uniqueness can be guaranteed. This fact allows us to define conjugate points and hence an analog of Jacobi's condition.

2 STATEMENT OF THE PROBLEM AND VARIATIONS AS DIFFERENTIALS

In this section we pose the simple fixed endpoint problem in the calculus of variations involving one delay in the phase coordinates. We introduce the first and second variations and show how, when a suitable norm is chosen on the space of trajectories, they can be defined as Fréchet differentials of the functional we are minimizing.

We are given the following:

i. An interval $[t_0,t_1]$ and a point ξ in R^n.

ii. A set A in Ψ, where Ψ denotes the usual topology of $[t_0,t_1] \times R^{3n}$.

iii. A positive number θ.

iv. A function ϕ mapping $[t_0-\theta,t_0]$ to R^n.

v. A function L mapping $[t_0,t_1] \times R^{3n}$ to R (the **Lagrangian**).

Denote the space of trajectories by:

$$X := \{x:[t_0-\theta,t_1] \to R^n \mid x \text{ is piecewise-}C^1\},$$

for all x in X set

$$\tilde{x}(t) := (t,x(t),x(t-\theta),\dot{x}(t)) \qquad t_0 \leqslant t \leqslant t_1$$

and let

$$X(A) := \{x \in X \mid \tilde{x}(t) \in A \ (t_0 \leqslant t \leqslant t_1) \text{ and } L \circ \tilde{x} \text{ is integrable}\}$$

(admissible trajectories)

and

$$X_\bullet(A) := \{x \in X(A) \mid x(t) = \phi(t) \quad \forall\ t \in [t_0-\theta,t_0] \text{ and } x(t_1) = \xi\}$$

(endpoint constraints).

The problem we shall be concerned with, which we label P(A), is that of minimizing the functional I(x) over $X_\bullet(A)$, where:

$$I(x):=\int_{t_0}^{t_1} L(\tilde{x}(t))dt \qquad \text{for all } x \in X(A).$$

Define **tubes** and **restricted tubes** about a trajectory x as follows: for all $\varepsilon > 0$, let

$$T_0(x;\varepsilon):=\{(t,y,u) \in [t_0,t_1] \times R^{2n} \mid |x(t) - y| < \varepsilon, |x(t-\theta) - u| < \varepsilon\}$$

and

$$T_1(x;\varepsilon):= \{(t,y,u,v) \in T_0(x;\varepsilon) \times R^n \mid |\dot{x}(t) - v| < \varepsilon\}.$$

We shall say that a trajectory x **solves P(A)** if

$$x \in S(A):=\{x \in X_\bullet(A) \mid I(x) \leqslant I(y) \quad \forall\, y \in X_\bullet(A)\},$$

x is a **strong minimum for P(A)** if there exists $\varepsilon > 0$ such that x solves $P((T_0(x;\varepsilon) \times R^n) \cap A)$ and x is a **weak minimum for P(A)** if there exists $\varepsilon > 0$ such that x solves $P(T_1(x;\varepsilon) \cap A)$. In other words, x is a strong (weak) minimum for P(A) if x in $X_\bullet(A)$ and there exists $\varepsilon > 0$ such that $I(y) \geqslant I(x)$ for all y in $X_\bullet(A)$ with $(t,y(t),y(t-\theta))$ in $T_0(x;\varepsilon)$ ($(t,y(t),y(t-\theta),\dot{y}(t))$ in $T_1(x;\varepsilon)$).

Let us denote the values of L by $L(t,x,u,\dot{x})$. Given a trajectory x in X, we define the **first variation of L** (with respect to x) by:

$$I'(x;y):=\int_{t_0}^{t_1} \{\langle L_x(\tilde{x}(t)) , y(t)\rangle + \langle L_u(\tilde{x}(t)) , y(t-\theta)\rangle + \langle L_{\dot{x}}(\tilde{x}(t)) , \dot{y}(t)\rangle\}dt$$

and the **second variation of L** (with respect to x) by:

$$I''(x;y):=\int_{t_0}^{t_1} 2\Omega(t,y(t),y(t-\theta),\dot{y}(t))dt$$

where, for all $(t,y,v,\dot{y})$ in $[t_0,t_1] \times R^{3n}$,

$$2\Omega(t,y,v,\dot{y}) := \langle y , L_{xx}(\tilde{x}(t))y\rangle + \langle v , L_{uu}(\tilde{x}(t))v\rangle + \langle \dot{y} , L_{\dot{x}\dot{x}}(\tilde{x}(t))\dot{y}\rangle$$

$$+ 2\{\langle y , L_{xu}(\tilde{x}(t))v\rangle + \langle y , L_{x\dot{x}}(\tilde{x}(t))\dot{y}\rangle + \langle \dot{y} , L_{\dot{x}u}(\tilde{x}(t))v\rangle\}.$$

The following proposition shows that the variations turn out to be differentials of I, if we introduce in X the norm

$$\| x\| := \sup\{|x(t)| + |\dot{x}(t)| \mid t_0-\theta \leq t \leq t_1\}.$$

1. Proposition: Assume L is $C^2(A;x,u,\dot{x})$. Then the functions $I'(x;\cdot)$ and $I''(x;\cdot)$ are the first and second Fréchet differentials at x of the functional I with respect to the space $(X,\|\ \|)$.

Proof: We begin by showing that X(A) is open as a subset of $(X,\|\ \|)$. For all x in X and $\varepsilon > 0$, let

$$B(x;\varepsilon) := \{y \in X \mid \|y - x\| < \varepsilon\}$$

and take x_0 in X(A). Since $\{\tilde{x}_0(t) \mid t \in [t_0,t_1]\}$ is compact in $[t_0,t_1] \times R^{3n}$ and A is open, there exists $\rho > 0$ such that the closure of $S := T_1(x_0;\rho)$ is contained in A. Let x be any trajectory in $B(x_0;\rho)$. So, for all t in $[t_0,t_1]$, $\tilde{x}(t) \in S \subset A$ and, consequently, x belongs to X(A), which proves that X(A) is open.

Now, since L and all its first and second partial derivatives with respect to x, u and $\dot{x}$ are continuous on A, they are uniformly continuous on the closure of S. So, if $\|x - x_0\| \to 0$, $L(\tilde{x}(t)) \to L(\tilde{x}_0(t))$ uniformly on $[t_0,t_1]$, and the same applies for the derivatives. Hence, for all $\varepsilon > 0$, there exists $\delta > 0$ such that, for all y in X and x in $B(x_0;\delta)$,

$$|I'(x;y) - I'(x_0;y)| < \varepsilon\|y\|$$

$$|I''(x;y) - I''(x_0;y)| < \varepsilon\|y\|^2$$

and so, if $\|x - x_0\| \to 0$, then $I'(x;y) \to I'(x_0;y)$ and $I''(x;y) \to I''(x_0;y)$ uniformly for $\|y\| \leqslant 1$.

Observe next that, if x is any trajectory in $B(x_0;\rho)$, so also is $x_0 + \lambda(x - x_0)$ for all λ in [0,1]. Applying Taylor's theorem to the function $\lambda \to I(x_0 + \lambda(x - x_0))$ we obtain, for all x in $B(x_0;\rho)$:

$$I(x) = I(x_0) + I'(x_0,x - x_0) + P_1(x_0,x - x_0) \tag{2.1}$$

where:

$$P_1(x_0,y) = \int_0^1 \{I'(x_0 + \lambda y;y) - I'(x_0;y)\}d\lambda$$

$$= \int_0^1 (1 - \lambda)I''(x_0 + \lambda y;y)d\lambda$$

and

$$I(x) = I(x_0) + I'(x_0,x - x_0) + \frac{1}{2} I''(x_0,x - x_0) + P_2(x_0,x - x_0)$$

where:

$$P_2(x_0,y) = \int_0^1 (1 - \lambda)\{I''(x_0 + \lambda y;y) - I''(x_0;y)\}d\lambda.$$

The required result now follows, since

$$\lim_{x \to x_0} \frac{P_1(x_0,x - x_0)}{\|x - x_0\|} = 0 \quad \text{and} \quad \lim_{x \to x_0} \frac{P_2(x_0,x - x_0)}{\|x - x_0\|^2} = 0.$$

It should be noted that this result was obtained by making use only of Taylor's theorem. It implies, in a natural way, sufficient conditions for a weak minimum. Let us denote by Y the set of trajectories vanishing at $[t_0-\theta,t_0]$ and t_1, i.e.,

$$Y := \{y \in X \mid y(t) = 0 \quad \forall\ t \in [t_0-\theta,t_0] \text{ and } y(t_1) = 0\}.$$

2. Corollary: Suppose $L \in C^2(A;x,u,\dot{x})$ and x_0 is a trajectory satisfying the endpoint constraints. If there exists $\varepsilon > 0$ such that $I'(x_0;y) = 0$ and $I''(x;y) > 0$ for all y in $Y - \{0\}$ and all x in $X_\bullet(A) \cap B(x_0;\varepsilon)$, then x_0 is a strict weak minimum for P(A).

Proof: Diminish ε so that (2.1) holds. Then

$$I(x) - I(x_0) = P_1(x_0, x - x_0) > 0$$

for all x in $X_\bullet(T_1(x_0;\varepsilon) \cap A) - \{x_0\}$ and the result follows.

3 NECESSARY AND SUFFICIENT CONDITIONS THROUGH VARIATIONS

In the preceding section we found sufficient conditions for a weak minimum by means of the concept of differentials without making use, at all, of the special form of the functions involved. The goal of the present section is to establish, through several properties particular of these functions, necessary and sufficient conditions both for weak and strong minima.

We begin by defining some sets of trajectories which are a natural extension of the ones we had for the delay free problem. For all A in $\mathcal{Y}$, let:

(1) Vanishing of the first variation in Y.

$$E(A) := \{x \in X(A) \mid I'(x;y) = 0 \quad \forall\, y \in Y\} \qquad \text{(set of extremals)}.$$

(2) Nonnegativity of the second variation in Y.

$$H(A) := \{x \in X(A) \mid I''(x;y) \geq 0 \quad \forall\, y \in Y\}.$$

(3) The condition of Weierstrass.

$$W(A) := \{x \in X(A) \mid E(t,x(t),x(t-\theta),\dot{x}(t),v) \geqslant 0 \quad \forall\ (t,x(t),x(t-\theta),v) \in A\}.$$

where the **Weierstrass 'excess function'** E, is given by:

$$E(t,x,u,\dot{x},v) := L(t,x,u,v) - L(t,x,u,\dot{x}) - \langle v - \dot{x}\ ,\ L_{\dot{x}}(t,x,u,\dot{x})\rangle.$$

(4) **The condition of Legendre.**

$$L(A) := \{x \in X(A) \mid L_{\dot{x}\dot{x}}(\tilde{x}(t)) \geqslant 0 \quad \forall\ t \in [t_0,t_1]\}.$$

The following theorem gives necessary conditions for a solution of P(A), where A is any set in Ψ. Thus, necessary conditions for a weak or strong minimum are derived if we replace A by $(T_1(x;\varepsilon) \cap A)$ or $(T_0(x;\varepsilon) \times R^n) \cap A$ respectively.

3. **Theorem:** Suppose L is continuous on A and x is a trajectory solving P(A). Then x lies in E(A), H(A), W(A) and L(A) whenever the derivatives of L involved (in each case) are continuous. In particular, if $L \in C^2(A;x,u,\dot{x})$, then:

$$S(A) \subset E(A) \cap H(A) \cap W(A) \cap L(A).$$

Proof: Each condition will be proved separately according to the smoothness required for the Lagrangian.

i. "$L \in C^1(A;x,u,\dot{x}) \rightarrow S(A) \subset E(A)$": Suppose x solves P(A) and y is any trajectory in Y. Since A is open, there exists $\delta > 0$ such that, for all $|\varepsilon| < \delta$, $x + \varepsilon y$ belongs to X(A). Define for all $|\varepsilon| < \delta$ and t in $[t_0,t_1]$,

$$F(t,\varepsilon) := L(t,x(t)+\varepsilon y(t),x(t-\theta)+\varepsilon y(t-\theta),\dot{x}(t)+\varepsilon\dot{y}(t))$$

$$f(\varepsilon) := \int_{t_0}^{t_1} F(t,\varepsilon)dt.$$

Since y belongs to Y, $x + \varepsilon y$ belongs to $X_\bullet(A)$ and thus $f(0) \leqslant f(\varepsilon)$ for all $|\varepsilon| < \delta$. By hypothesis, L, L_x, L_u and $L_{\dot{x}}$ are continuous on A. This implies that F and F_ε are piecewise continuous, and so:

$$0 = f'(0) = \int_{t_0}^{t_1} F_\varepsilon(t,0)dt = I'(x;y).$$

ii. "$L \in C^2(A;x,u,\dot{x}) \Rightarrow S(A) \subset H(A)$": An analogous argument to that of (i) applies, and we obtain:

$$0 \leqslant f''(0) = \int_{t_0}^{t_1} F_{\varepsilon\varepsilon}(t,0)dt = I''(x;y).$$

iii. "$L \in C^2(A;\dot{x}) \Rightarrow W(A) \subset L(A)$": Let $x \in W(A)$, fix $t \in [t_0,t_1]$ and define, for all v in R^n:

$$G(v) := E(t,x(t),x(t-\theta),\dot{x}(t),v).$$

By definition of the excess function, G has a local minimum at $\dot{x}(t)$. So, for all c in R^n,

$$0 \leqslant \langle c \, , \, G''(\dot{x}(t))c\rangle = \langle c \, , \, L_{\dot{x}\dot{x}}(\tilde{x}(t))c\rangle$$

which implies that $x \in L(A)$.

iv. "$L \in C^1(A;\dot{x}) \Rightarrow S(A) \subset W(A)$": Suppose x_0 is a trajectory in S(A). Consider any point s in (t_0,t_1) at which $\dot{x}_0$ is continuous and let v in R^n be such that $(s,x(s),x(s-\theta),v)$ belongs to A. Now, choose any μ in $(0,t_1-s)$ and define the following family of trajectories, for all δ in $[0,\mu]$ and ε in $[0,1)$:

$$x(t;\varepsilon,\delta) := x_0(t) \qquad t \in [t_0-\theta,s] \cup [s+\delta,t_1]$$

$$x(t;\varepsilon,\delta) := x_0(t) + (t-s)(v-\dot{x}_0(s)) \qquad t \in [s, s+\varepsilon\delta]$$

$$x(t;\varepsilon,\delta) := x_0(t) + \lambda(\varepsilon)(s+\delta-t)(v-\dot{x}_0(s)) \qquad t \in [s+\varepsilon\delta, s+\delta]$$

where $\lambda(\varepsilon) = \varepsilon/(1-\varepsilon)$.

Since A is open, there exists $\eta > 0$ such that, for all δ in $[0,\mu]$ and ε in $[0,\eta]$, $x(\cdot;\varepsilon,\delta)$ belongs to $X_\bullet(A)$. Therefore,

$$0 \leqslant I(x(\cdot;\varepsilon,\delta)) - I(x_0) = F(\varepsilon,\delta) + G(\varepsilon,\delta), \tag{2.2}$$

where:

$$F(\varepsilon,\delta) := \int_s^{s+\varepsilon\delta} \{L(t, x(t;\varepsilon,\delta), x(t-\theta;\varepsilon,\delta), \dot{x}_0(t)+v-\dot{x}_0(s)) - L(\tilde{x}_0(t))\}dt$$

$$G(\varepsilon,\delta) := \int_{s+\varepsilon\delta}^{s+\delta} \{L(t, x(t;\varepsilon,\delta), x(t-\theta;\varepsilon,\delta), \dot{x}_0(t)-\lambda(\varepsilon)(v-\dot{x}_0(s))) - L(\tilde{x}_0(t))\}dt.$$

Now, since L is continuous on A, for all $0 < |\varepsilon| \leqslant \eta$,

$$\lim_{\delta \to 0} \frac{F(\varepsilon,\delta)}{\varepsilon\delta} = L(s, x_0(s), x_0(s-\theta), v) - L(\tilde{x}_0(s))$$

and

$$\lim_{\delta \to 0} \frac{G(\varepsilon,\delta)}{\varepsilon\delta} = \frac{1}{\lambda(\varepsilon)}[L(s, x_0(s), x_0(s-\theta), \dot{x}_0(s) - \lambda(\varepsilon)(v - \dot{x}_0(s)))) - L(\tilde{x}_0(s))].$$

In view of (2.2), we have:

$$0 \leqslant \lim_{\varepsilon \to 0} \left[\lim_{\delta \to 0} \frac{F(\varepsilon,\delta)}{\varepsilon\delta} + \lim_{\delta \to 0} \frac{G(\varepsilon,\delta)}{\varepsilon\delta} \right]$$

$$= E(s, x_0(s), x_0(s-\theta), \dot{x}_0(t), v).$$

Finally, by continuity of L and $L_{\dot{x}}$, the condition of Weierstrass also holds at the end and corner points of x_0, and the result follows.

As in the delay free case, slightly strengthening the previous sets we obtain sufficient conditions. Let us consider the following definitions:

(1)′ **Nonsingular extremals.**

$$E'(A) := \{x \ \varepsilon \ E(A) \ | \ |L_{\dot{x}\dot{x}}(\tilde{x}(t))| \neq 0 \quad \forall \ t \ \varepsilon \ [t_0, t_1]\}$$

(2)′ **Positivity of the second variation in Y − { 0 } .**

$$H'(A) := \{x \ \varepsilon \ H(A) \ | \ I''(x;y) > 0 \quad \forall \ y \ \varepsilon \ Y - \{0\}\}.$$

(3)′ **Strengthened Weierstrass condition.**

$$W(A;\varepsilon) := \{x_0 \ \varepsilon \ W(A) \ | \ E(t,x,u,\dot{x},v) \geqslant 0 \quad \forall \ (t,x,u,\dot{x}) \ \varepsilon \ T_1(x_0;\varepsilon) \text{ and } (t,x,u,v) \ \varepsilon \ A\}.$$

(4)′ **Strengthened Legendre condition.**

$$L'(A) := \{x \ \varepsilon \ L(A) \ | \ L_{\dot{x}\dot{x}}(\tilde{x}(t)) > 0 \quad \forall \ t \ \varepsilon \ [t_0, t_1]\}.$$

We shall find it useful to first establish the following property of the Weierstrass function.

4. **Lemma:** Suppose L is $C^2(A;\dot{x})$ and x_0 is a nonsingular trajectory that belongs to $W(A;\varepsilon)$ for some $\varepsilon > 0$. Then there exist δ and $h > 0$ such that, for all $(t,x,u,\dot{x})$ in $T_1(x_0;\delta)$ and (t,x,u,v) in A,

$$E(t,x,u,\dot{x},v) \geqslant h \, [(1 + |v - \dot{x}|^2)^{1/2} - 1]. \tag{2.3}$$

Proof: By Theorem 2(iii), x_0 belongs to L(A) and, since x_0 is nonsingular, it also belongs to L′(A). So, by continuity of $L_{\dot{x}\dot{x}}$, there exist ε_0 and $h_0 > 0$ such that, for all c in R^n and $(t,x,u,\dot{x})$ in $T_1(x_0;\varepsilon_0)$, $\langle c \, , \, L_{\dot{x}\dot{x}}(t,x,u,\dot{x})c\rangle \geqslant h_0|c|^2$. Without loss of generality, $\varepsilon_0 < \varepsilon$. Now, by Taylor's theorem, for all $(t,x,u,\dot{x})$ in $T_1(x_0;\varepsilon_0)$ and (t,x,u,v) in $T_1(x_0;\varepsilon_0)$,

$$E(t,x,u,\dot{x},v) = L(t,x,u,v) - L(t,x,u,\dot{x}) - \langle v - \dot{x} \, , \, L_{\dot{x}}(t,x,u,\dot{x})\rangle$$

$$= \int_0^1 (1 - \lambda)(\langle v - \dot{x} , L_{\dot{x}\dot{x}}(t,x,u,\dot{x}+\lambda(v-\dot{x}))(v - \dot{x})\rangle)d\lambda$$

$$\geqslant \frac{h_0}{2} (|v - \dot{x}|^2).$$

Let $\delta > 0$ be such that the closure of $T_1(x_0;\delta)$ is contained in $T_1(x_0;\varepsilon)$ and $\rho > 0$ such that, for all $|c| \leqslant \rho$ and $(t,x,u,\dot{x})$ in $T_1(x_0;\delta)$, $(t,x,u,\dot{x} + c)$ belongs to $T_1(x_0;\varepsilon_0)$.

Now, take $(t,x,u,\dot{x})$ in $T_1(x_0;\delta)$ and (t,x,u,v) in A. Observe that, if (t,x,u,v) is in $T_1(x_0;\varepsilon_0)$, then

$$E(t,x,u,\dot{x},v) \geqslant \frac{h_0}{2} (|v - \dot{x}|^2) \geqslant h_0 [(1 + |v - \dot{x}|^2)^{1/2} - 1]$$

and so (2.3) holds with $h = h_0$. For the case $(t,x,u,v) \notin T_1(x_0;\varepsilon_0)$, let

$$k:= \frac{|v - \dot{x}|}{\rho} \quad \text{and} \quad c:= \frac{(v - \dot{x})}{k}$$

and observe that $|c| = \rho$ and $k > 1$. Hence,

$$\begin{aligned} E(t,x,u,\dot{x},v) &= E(t,x,u,\dot{x},\dot{x} + kc) \\ &= E(t,x,u,\dot{x} + c,\dot{x} + kc) + kE(t,x,u,\dot{x},\dot{x} + c) \\ &\qquad + (k - 1)E(t,x,u,\dot{x} + c,\dot{x}) \\ &\geqslant k\, E(t,x,u,\dot{x},\dot{x} + c) \geqslant \frac{kh_0}{2} |c|^2 \\ &= \frac{h_0}{2} |c||kc| \geqslant \frac{h_0\rho}{2} [(1 + |kc|^2)^{1/2} - 1] \\ &= \frac{h_0\rho}{2} [(1 + |v - \dot{x}|^2)^{1/2} - 1]. \end{aligned}$$

Therefore, equation (2.3) holds with $h = \min\{h_0,(h_0\rho)/2\}$ and the result follows.

We are going to prove now the main result for sufficient conditions in terms of the first and second variations. What we shall prove is actually stronger than the statement of the theorem. The conditions will turn out to be sufficient for a local minimum, not only with respect to the space X of piecewise-C^1 trajectories, but also with respect to the class X′ of all absolutely continuous functions mapping $[t_0-\theta, t_1]$ to R^n. Throughout the proof, X′(A) and $X_\bullet'(A)$ will denote, respectively, the sets of functions in X′ being admissible and satisfying the endpoint constraints. The main ideas that follow are based on the proof for the delay free case given by M. R. Hestenes (see [18]).

5. **Theorem:** Suppose $L \in C^2(A;x,u,\dot{x})$ and x_0 is a trajectory satisfying the endpoint constraints.

i. If x_0 lies in E(A), H′(A) and L′(A), then x_0 is a (strict) weak minimum for P(A) of class $C^1([t_0,t_1])$.

ii. If, for some $\varepsilon > 0$, x_0 lies in E(A), H′(A), $W(A;\varepsilon)$ and L′(A), then x_0 is a (strict) strong minimum for P(A) of class $C^1([t_0,t_1])$.

Proof: In Corollary 8 (which is independent of this theorem), we shall show that, if x_0 is a trajectory satisfying the assumptions of (i), then x_0 belongs to $W(T_1(x_0;\varepsilon);\varepsilon)$ for some $\varepsilon > 0$. If we suppose that assertion (ii) is true, x_0 will be a strong minimum for $P(T_1(x_0;\varepsilon))$, i.e., there will exist $\delta > 0$ such that $P((T_0(x_0;\delta) \cap R^n) \cap (T_1(x_0;\varepsilon)))$ is solved by x_0. Setting $\mu := \min\{\varepsilon,\delta\}$, it follows that:

$$T_1(x_0;\mu) \cap A \subset T_1(x_0;\mu) \subset (T_0(x_0;\delta) \cap R^n) \cap T_1(x_0;\varepsilon).$$

But this implies that x_0 is a weak minimum for P(A) and so (i) holds. The smoothness of the trajectory is also a consequence of Corollary 8, and so it remains only to prove that, if for some $\varepsilon > 0$

$$x_0 \in X_\bullet(A) \cap E'(A) \cap H'(A) \cap W(A;\varepsilon), \tag{2.4}$$

then x_0 is a strong minimum for P(A).

Let us start considering the following strengthened set of solutions:

$$S(A;\rho) := \{x \in S(A) \mid I(y) \geqslant I(x) + \rho D(y - x) \quad \forall\, y \in X_\bullet'(A)\}$$

where, for all y in X' and c in $\mathbf{R}^n$,

$$D(y) := \int_{t_0}^{t_1} f(\dot{y}(t))dt \quad \text{and} \quad f(c) := (1 + |c|^2)^{1/2} - 1.$$

The proof consists in showing that the conditions imposed on L and x_0, imply the existence of some ρ and $\delta > 0$ such that x_0 belongs to $S(T_0(x_0;\rho) \cap A;\delta)$, which in turn implies that x_0 is a strict strong minimum for P(A). Suppose the contrary, that is,

$$x_0 \notin \bigcup_{\rho,\delta>0} S(T_0(x_0;\rho) \cap A;\delta). \tag{2.5}$$

We shall show that $x_0 \notin H'(A)$ thus contradicting the assumption (2.4).

To begin with, observe that, for all x in $X'(A)$,

$$I(x) = I(x_0) + I'(x_0;x - x_0) + K(x) + E^*(x)$$

where:

$$E^*(x) := \int_{t_0}^{t_1} E(t,x(t),x(t-\theta),\dot{x}_0(t),\dot{x}(t))dt$$

$$K(x) := \int_{t_0}^{t_1} \{M(t,x(t),x(t-\theta)) + \langle \dot{x}(t) - \dot{x}_0(t) \;,\; N(t,x(t),x(t-\theta))\rangle\}dt$$

and

$$N(t,y,u) := L_{\dot{x}}(t,y,u,\dot{x}_0(t)) - L_{\dot{x}}(\tilde{x}_0(t))$$

$$M(t,y,u) := L(t,y,u,\dot{x}_0(t)) - L(\tilde{x}_0(t)) - \langle L_x(\tilde{x}_0(t)) \;,\; y - x_0(t)\rangle - \langle L_u(\tilde{x}_0(t)) \;,\; u - x_0(t-\theta)\rangle.$$

Moreover, by Taylor's theorem, there exists $\mu > 0$ such that, for all

(t,y,u) in $T_0(x_0;\mu)$,

$$M(t,y,u) = \frac{1}{2} \langle y - x_0(t) \ , \ P_1(t,y,u)(y - x_0(t))\rangle$$
$$+ \frac{1}{2} \langle u - x_0(t-\theta) \ , \ P_2(t,y,u)(u - x_0(t-\theta))\rangle$$
$$+ \langle y - x_0(t) \ , \ P_3(t,y,u)(u - x_0(t-\theta))\rangle$$

and

$$N(t,y,u) = Q_1(t,y,u)(y - x_0(t)) + Q_2(t,y,u)(u - x_0(t-\theta)),$$

where:

$$P_1(t,y,u) := 2 \int_0^1 (1 - \lambda)L_{xx}d\lambda, \quad P_2(t,y,u) := 2 \int_0^1 (1 - \lambda)L_{uu}d\lambda,$$

$$P_3(t,y,u) := 2 \int_0^1 (1 - \lambda)L_{xu}d\lambda, \quad Q_1(t,y,u) := \int_0^1 L_{\dot{x}x}d\lambda, \quad Q_2(t,y,u) := \int_0^1 L_{\dot{x}u}d\lambda$$

with all derivatives of L evaluated at

$$(t,x_0(t) + \lambda(y - x_0(t)),x_0(t-\theta) + \lambda(u - x_0(t-\theta)),\dot{x}_0(t)).$$

Next, we show that there exist δ_0, α and $h > 0$ such that, for all x in $X'(T_0(x_0;\delta_0) \cap A)$,

$$E^*(x) \geqslant hD(x - x_0) \tag{2.6}$$

and

$$|K(x)| \leqslant \alpha\| x - x_0\| (1 + D(x - x_0)) \tag{2.7}$$

where, in contrast with the norm used in Section 2,

$$\| x\| := \sup\{|x(t)| \mid t \in [t_0,t_1]\}.$$

To prove it, let δ and $h > 0$ satisfy the conclusion of Lemma 4, and let x be any trajectory in $X'(T_0(x_0;\delta) \cap A)$. In view of (2.3), we have

$$E^*(x) = \int_{t_0}^{t_1} E(t,x(t),x(t-\theta),\dot{x}_0(t),\dot{x}(t))dt$$

$$\geq h \int_{t_0}^{t_1} f(\dot{x}(t) - \dot{x}_0(t))dt = hD(x - x_0)$$

and (2.6) holds. For (2.7), we can clearly choose $\alpha' > 0$ such that, for all x in $X'(T_0(x_0;\mu) \cap A)$ and t in $[t_0,t_1]$,

$$| M(t,x(t),x(t-\theta)) + \langle \dot{x}(t) - \dot{x}_0(t) , N(t,x(t),x(t-\theta))\rangle |$$

$$\leq \alpha'[|x(t) - x_0(t)| + |x(t-\theta) - x_0(t-\theta)|](1 + |\dot{x}(t) - \dot{x}_0(t)|^2)^{1/2}.$$

But this implies that:

$$|K(x)| \leq 2\alpha' \| x - x_0 \| \int_{t_0}^{t_1} (1 + f(\dot{x}(t) - \dot{x}_0(t)))dt$$

$$\leq \alpha \| x - x_0 \| (1 + D(x - x_0))$$

with $\alpha := 2\max\{\alpha',\alpha'(t_1 - t_0)\}$. Hence, (2.6) and (2.7) hold with $0 < \delta_0 < \min\{\mu,\delta\}$.

Now, in view of (2.5), for every natural number q there exists x_q in $X_\bullet'(A)$ such that:

$$\| x_q - x_0 \| < \min\{\delta_0,1/q\} \text{ and } I(x_q) - I(x_0) < D(x_q - x_0)/q. \qquad (2.8)$$

An important property we shall be using of the sequence $\{x_q\}$, is that it contains a subsequence (we do not relabel), such that

$$\dot{x}_q(t) \to \dot{x}_0(t), \quad q \to \infty \text{ almost uniformly on } [t_0,t_1]. \qquad (2.9)$$

This follows observing first that, since $x_q - x_0 \in Y$ and $x_0 \in E(A)$,

$$I(x_q) - I(x_0) = K(x_q) + E^*(x_q)$$

$$\geqslant - \alpha\| x_q - x_0\| + D(x_q - x_0)(h - \alpha\| x_q - x_0\|)$$

which implies that, for all q in **N**,

$$D(x_q - x_0)(h - 1/q - \alpha/q) < \alpha/q$$

and so $D(x_q - x_0) \to 0, \quad q \to \infty$. Setting

$$w_q(t) := [1 + \frac{1}{2} f(\dot{x}_q(t) - \dot{x}_0(t))]^{1/2},$$

observe that, by the inequality of Schwarz,

$$\left| \int_{t_0}^{t_1} |\dot{x}_q(t) - \dot{x}_0(t)| dt \right|^2 \leqslant \int_{t_0}^{t_1} \frac{|\dot{x}_q(t) - \dot{x}_0(t)|^2 dt}{w_q(t)^2} \int_{t_0}^{t_1} w_q(t)^2 dt$$

$$= D(x_q - x_0)(2(t_1 - t_0) + D(x_q - x_0)).$$

So, $\dot{x}_q \to \dot{x}_0$ in the L^1-norm and (2.9) holds.

Now, in view of (2.8), $d_q := (2D(x_q - x_0))^{1/2}$ is strictly positive for all q in **N**, and we can define the following sequence of absolutely continuous functions:

$$y_q(t) := (x_q(t) - x_0(t))/d_q \qquad t_0 - \theta \leqslant t \leqslant t_1, \quad q \in \mathbf{N}.$$

Observe that, for all y in X′, the functional D satisfies

$$D(y) = \int_{t_0}^{t_1} \frac{|\dot{y}_q(t)|^2 dt}{2 + f(\dot{y}(t))} \tag{2.10}$$

and so, for all q in **N**,

$$\int_{t_0}^{t_1} \frac{|\dot{y}_q(t)|^2 dt}{w_q(t)^2} = 1. \tag{2.11}$$

Consequently, $\{\dot{y}_q/w_q\}$ is a sequence of square-integrable functions and, from the last relation, there exist a subsequence of $\{y_q\}$ and a function

u_0 in $L^2([t_0,t_1])$, such that $\dot{y}_q/w_q$ converges weakly (on $L^2([t_0,t_1])$) to u_0, that is,

$$\lim_{q\to\infty}\int_{t_0}^{t_1}\left\langle g(t)\ ,\ \frac{\dot{y}_q(t)}{w_q(t)}\right\rangle dt=\int_{t_0}^{t_1}\langle g(t)\ ,\ u_0(t)\rangle dt \tag{2.12}$$

for every square-integrable function $g(t)$ on $[t_0,t_1]$.

Define the following function, which belongs to X' and has a square-integrable derivative:

$$y_0(t):=\begin{cases}0 & t_0-\theta\leqslant t\leqslant t_0\\ \int_{t_0}^{t}u_0(s)ds & t_0\leqslant t\leqslant t_1\end{cases}$$

We wish to show that y_0 satisfies the following:

i. y_0 belongs to Y.

ii. $I''(x_0;y_0)\leqslant 0$.

iii. y_0 is not identically zero on $[t_0,t_1]$.

This will imply that $x_0 \notin H'(A)$ and the proof will be complete. For this purpose, we prove next a few convergence properties of the sequence $\{y_q\}$.

To start with, observe that, for all q in $\mathbf{N}$,

$$\int_{t_0}^{t_1}|w_q(t)-1|^2dt=\int_{t_0}^{t_1}(w_q(t)^2-1)dt-2\int_{t_0}^{t_1}(w_q(t)-1)dt$$

and, since $w_q(t)^2\geqslant w_q(t)\geqslant 1$ almost everywhere, we also have that

$$0\leqslant\int_{t_0}^{t_1}(w_q(t)-1)dt\leqslant\int_{t_0}^{t_1}(w_q(t)^2-1)dt\leqslant D(x_q-x_0).$$

These facts clearly imply that

$$\lim_{q\to\infty}\int_{t_0}^{t_1}(w_q(t)^2-1)dt = \lim_{q\to\infty}\int_{t_0}^{t_1}(w_q(t)-1)dt$$

$$= \lim_{q\to\infty}\int_{t_0}^{t_1}|w_q(t)-1|^2dt = 0 \tag{2.13}$$

With the help of these equalities, let us prove that $\dot{y}_q$ converges weakly on L^1 to $\dot{y}_0$. Let g be any function in L^∞ and note that, for all q in **N**,

$$\left|\int_{t_0}^{t_1} \langle g(t)(w_q(t) - 1) , \frac{\dot{y}_q(t)}{w_q(t)} \rangle dt\right|^2 \leqslant \int_{t_0}^{t_1} |g(t)|^2|w_q(t) - 1|^2dt$$

and

$$\langle g(t) , \dot{y}_q(t)\rangle = \langle g(t) , \frac{\dot{y}_q(t)}{w_q(t)} \rangle + \langle g(t)(w_q(t) - 1) , \frac{\dot{y}_q(t)}{w_q(t)} \rangle.$$

Using (2.12), (2.13) and the boundedness of g, we have

$$\lim_{q\to\infty}\int_{t_0}^{t_1} \langle g(t) , \dot{y}_q(t)\rangle dt = \lim_{q\to\infty}\int_{t_0}^{t_1} \langle g(t) , \frac{\dot{y}_q(t)}{w_q(t)} \rangle dt$$

$$= \int_{t_0}^{t_1} \langle g(t) , \dot{y}_0(t)\rangle dt \tag{2.14}$$

and the weak convergence follows. Note that (2.14) holds also for all g in L^2 integrating over a measurable set where $w_q(t)$ tends to 1 uniformly. Moreover, if R_q and R are continuous functions such that $R_q(t) \to R(t)$ uniformly on $[t_0,t_1]$, then

$$\lim_{q\to\infty}\int_{t_0}^{t_1} \langle R_q(t) , \dot{y}_q(t)\rangle dt = \int_{t_0}^{t_1} \langle R(t) , \dot{y}_0(t)\rangle dt. \tag{2.15}$$

This last assertion follows observing that, by (2.13) and (2.16) below, there exists a constant $T > 0$ such that, for all q in **N**,

$$\left|\int_{t_0}^{t_1} |\dot{y}_q(t)|dt\right|^2 \leqslant \int_{t_0}^{t_1} w_q(t)^2 dt \leqslant T$$

and so, in view of (2.14),

$$\lim_{q\to\infty}\int_{t_0}^{t_1} \langle R_q(t) , \dot{y}_q(t)\rangle dt = \lim_{q\to\infty}\int_{t_0}^{t_1} \langle R(t) , \dot{y}_q(t)\rangle dt = \int_{t_0}^{t_1} \langle R(t) , \dot{y}_0(t)\rangle dt$$

and (2.15) holds.

Next, we prove that y_q converges to y_0 uniformly on $[t_0,t_1]$. To this end, observe first that, by virtue of the weak convergence of $\dot{y}_q$ to $\dot{y}_0$ in L^1, for all t in $[t_0,t_1]$ we have

$$\lim_{q\to\infty}\int_{t_0}^{t} \dot{y}_q(s)ds = \lim_{q\to\infty}\int_{t_0}^{t} \dot{y}_0(s)ds$$

and so y_q converges pointwise to y_0. Now, let S be any measurable set in $[t_0,t_1]$. By the inequality of Schwarz and (2.11),

$$\left|\int_S \dot{y}_q(t)dt\right|^2 \leqslant \left(\int_S \frac{|\dot{y}_q(t)|^2 dt}{w_q(t)^2}\right)\left(\int_S w_q(t)^2 dt\right)$$

$$\leqslant \int_S w_q(t)^2 dt$$

$$= m(s) + \int_S (w_q(t)^2 - 1)dt \qquad (2.16)$$

where m(S) is the measure of S. Given $\varepsilon > 0$, let q_0 in **N** be such that

$$q \geqslant q_0 \Rightarrow \int_{t_0}^{t_1} (w_q(t)^2 - 1)dt < \varepsilon/2,$$

and $0 < \delta < \varepsilon/2$ such that, for all $q < q_0$,

$$m(S) < \delta \Rightarrow \left|\int_S \dot{y}_q(t)dt\right| < \varepsilon$$

So, if $m(S) < \varepsilon/2$, we have that, for all q in **N**, $|\int_S \dot{y}_q(t)dt| < \varepsilon$. Consequently, the sequence of integrals $\{\int_{t_0}^{t} \dot{y}_q(t)dt\}$ and hence also $\{y_q\}$, are equi-absolutely continuous on $[t_0,t_1]$ and so $y_q \to y_0$ uniformly on $[t_0,t_1]$.

We shall need one last result concerning the Weierstrass function, namely,

$$\liminf_{q \to \infty} \frac{E^*(x_q)}{d_q^{\,2}} \geq \frac{1}{2} \int_{t_0}^{t_1} \langle \dot{y}_0(t) \ , \ L_{\dot{x}\dot{x}}(\tilde{x}_0(t))\dot{y}_0(t)\rangle dt. \tag{2.17}$$

By virtue of (2.9), it suffices to prove it on $S \subset [t_0,t_1]$ where $\dot{x}_q(t) \to \dot{x}_0(t)$ uniformly. By Taylor's theorem, there exists q_0 in $\mathbf{N}$ such that, for all $q \geq q_0$ and all t in S,

$$\frac{1}{d_q^{\,2}} [E(t,x_q(t),x_q(t-\theta),\dot{x}_0(t),\dot{x}_q(t))] = \frac{1}{2} \langle \dot{y}_q(t) \ , \ R_q(t)\dot{y}_q(t)\rangle$$

where:

$$R_q(t) := 2 \int_0^1 (1 - \lambda)[L_{\dot{x}\dot{x}}(t,x_q(t),x_q(t-\theta),\dot{x}_0(t) + \lambda(\dot{x}_q(t)-\dot{x}_0(t)))]d\lambda.$$

Let $R(t) := L_{\dot{x}\dot{x}}(\tilde{x}_0(t))$ and observe that, for all t in S,

$$|\langle \dot{y}_q(t) \ , \ (R_q(t) - R(t))\dot{y}_q(t)\rangle| \leq \| R_q(t) - R(t)\| \ |\dot{y}_q(t)|^2$$
$$\leq M_q \frac{|\dot{y}_q(t)|^2}{w_q(t)^2}$$

where:

$$M_q := \sup\{[\| R_q(t) - R(t))\|^2 \ w_q(t)^4]^{1/2} \mid t \in S\}.$$

Since $R_q(t) \to R(t)$ and $w_q(t) \to 1$, both uniformly on S, $M_q \to 0$ as $q \to \infty$. Observe also that

$$\langle \dot{y}_q(t) \ , \ R(t)\dot{y}_q(t)\rangle = \langle \dot{y}_0(t) \ , \ R(t)\dot{y}_0(t)\rangle + 2\langle \dot{y}_q(t) - \dot{y}_0(t) \ , \ R(t)\dot{y}_0(t)\rangle$$
$$+ \langle \dot{y}_q(t) - \dot{y}_0(t) \ , \ R(t)(\dot{y}_q(t) - \dot{y}_0(t))\rangle.$$

As we mentioned before, (2.14) holds replacing $g(t)$ by $R(t)\dot{y}_0(t)$ and integrating over S. Consequently,

$$\liminf_{q\to\infty} \int_S \langle \dot{y}_q(t)\ ,\ R_q(t)\dot{y}_q(t)\rangle dt$$

$$= \liminf_{q\to\infty} \int_S \langle \dot{y}_q(t)\ ,\ R(t)\dot{y}_q(t)\rangle dt$$

$$= \int_S \langle \dot{y}_0(t)\ ,\ R(t)\dot{y}_0(t)\rangle dt$$

$$+ \liminf_{q\to\infty} \int_S \langle \dot{y}_q(t) - \dot{y}_0(t)\ ,\ R(t)(\dot{y}_q(t) - \dot{y}_0(t))\rangle dt$$

and finally, since $R(t)$ is positive definite, (2.17) follows.

With the help of these results we are now in a position to prove that y_0 satisfies (i)-(iii). From their definition, each y_q belongs to Y and, since y_q converges to y_0 uniformly on $[t_0,t_1]$, y_0 belongs to Y (for this result it clearly suffices the pointwise convergence). So (i) holds. For (ii), observe that, by definition of the functional K,

$$\frac{K(x_q)}{d_q^{\,2}} = \frac{1}{2}\int_{t_0}^{t_1} \{\ \langle y_q(t)\ ,\ P_1\, y_q(t)\rangle + \langle y_q(t-\theta)\ ,\ P_2\, y_q(t-\theta)\rangle$$

$$+\ 2\ \langle y_q(t)\ ,\ P_3\, y_q(t-\theta)\rangle + 2\ \langle \dot{y}_q(t)\ ,\ Q_1\, y_q(t) + Q_2\, y_q(t-\theta)\rangle\ \}dt$$

where P_i and Q_i are evaluated at $(t,x_q(t),x_q(t-\theta))$ and q is sufficiently large. So we have:

$$\frac{1}{2}\, I''(x_0;y_0) = \lim_{q\to\infty} \frac{K(x_q)}{d_q^{\,2}} + \frac{1}{2}\int_{t_0}^{t_1} \langle \dot{y}_0(t)\ ,\ R(t))\dot{y}_0(t)\rangle dt] \qquad \text{(by (2.15))}$$

$$\leqslant \lim_{q\to\infty} \frac{K(x_q)}{d_q^{\,2}} + \liminf_{q\to\infty} \frac{E^*(x_q)}{d_q^{\,2}} \qquad \text{(by (2.17))}$$

$$= \liminf_{q\to\infty} \frac{I(x_q) - I(x_0)}{d_q^{\,2}} \leqslant 0 \qquad \text{(by (2.8)).}$$

Consequently, (ii) holds. Finally, for (iii), note that, if $y_0 \equiv 0$ then $\lim_{q \to \infty} K(x_q)/(d_q^2) = 0$ and so, from the last relation and (2.6),

$$\frac{h}{2} \leqslant \liminf_{q \to \infty} \frac{E^*(x_q)}{d_q^2} \leqslant 0$$

contradicting the positivity of h. The proof is now complete.

4 THE FIRST VARIATION: EULER'S EQUATION

We start characterizing the set E(A). Both Euler's equation and its consequences related to the smoothness of extremals can easily be extended to the delay problem.

6. **Proposition:** Suppose x is a trajectory in X(A). Then x is an extremal if, and only if, there exists a constant c in R^n such that, for all t in $[t_0,t_1]$,

$$L_{\dot{x}}(\tilde{x}(t)) = P(t;x) + c$$

where $P(\cdot;x)$ is given by:

$$P(t;x) := \int_{t_0}^{t} L_x(\tilde{x}(s))ds + \int_{t_0}^{t+\theta} L_u(\tilde{x}(s))ds \quad t_0 \leqslant t \leqslant t_1-\theta$$

$$:= \int_{t_0}^{t} L_x(\tilde{x}(s))ds + \int_{t_0}^{t_1} L_u(\tilde{x}(s))ds \quad t_1-\theta \leqslant t \leqslant t_1.$$

Proof: By definition of the first variation and the function $P(\cdot;x)$, observe that, for all x and y in X,

$$I'(x;y) = \int_{t_0}^{t_1} \{\langle y(t) , \dot{P}(t;x)\rangle + \langle\dot{y}(t) , L_{\dot{x}}(\tilde{x}(t))\rangle\}dt$$

$$+ \int_{t_0-\theta}^{t_0} \langle y(t) , L_u(\tilde{x}(t+\theta))\rangle dt.$$

The "if" part follows directly from this relation since, given y in Y,

$$0 = \int_{t_0}^{t_1} \frac{d}{dt}\langle y(t) \, , \, P(t;x) + c\rangle dt$$

$$= \int_{t_0}^{t_1} \{\langle y(t) \, , \, \dot{P}(t;x)\rangle + \langle \dot{y}(t) \, , \, L_{\dot{x}}(\tilde{x}(t))\rangle\}dt = I'(x;y).$$

Hence, we assume x in X(A) is such that $I'(x;y) = 0$ for all y in Y. Define:

$$c := \frac{1}{t_1 - t_0} \int_{t_0}^{t_1} (L_{\dot{x}}(\tilde{x}(t)) - P(t;x))dt$$

and

$$z(t) := \begin{cases} 0 & t_0-\theta \leqslant t \leqslant t_0 \\ \int_{t_0}^{t} (L_{\dot{x}}(\tilde{x}(s)) - P(s;x) - c)ds & t_0 \leqslant t \leqslant t_1 \end{cases}$$

so that $z \in Y$ and $\dot{z}(t) = L_{\dot{x}}(\tilde{x}(t)) - P(t;x) - c$. Clearly,

$$0 = \int_{t_0}^{t_1} \frac{d}{dt}\langle z(t) \, , \, P(t;x) + c\rangle dt$$

$$= \int_{t_0}^{t_1} \{\langle z(t) \, , \, \dot{P}(t;x)\rangle + \langle \dot{z}(t) \, , \, L_{\dot{x}}(\tilde{x}(t)) - \dot{z}(t)\rangle\}dt$$

$$= I'(x;z) - \int_{t_0}^{t_1} |\dot{z}(t)|^2 dt = - \int_{t_0}^{t_1} |\dot{z}(t)|^2 dt.$$

Consequently, $\dot{z}(t) = 0$ for all t in $[t_0, t_1]$ and the result follows.

7. **Corollary:** For any set A in Ψ, the following holds:

i. **Weierstrass-Erdmann corner condition.** If $x \in E(A)$ then $L_{\dot{x}} \circ \tilde{x}$ is continuous on $[t_0,t_1]$.

ii. Hilbert Differentiability Theorem. Assume L_x and L_u are $C^{r-2}(A)$, $L_{\dot{x}}$ is $C^{r-1}(A)$ $(r \geqslant 2)$, and $x_0 \in E'(A) \cap C^1([t_0,t_1])$. Then:

a. x_0 is C^r on $[t_0,t_1-\theta]$ and C^r on $[t_1-\theta,t_0]$.

b. If $L_u(\tilde{x}_0(t))$ vanishes at t_1, together with all derivatives up to order r-1, then x_0 is C^r on $[t_0,t_1]$.

iii. Regularity. If L is **regular** on A (A is convex in $\dot{x}$ and $L_{\dot{x}\dot{x}}(t,x,u,\dot{x}) > 0$ for all $(t,x,u,\dot{x})$ in A) and $L_{\dot{x}\dot{x}}$ is continuous on A, then:

a. $E(t,x,u,\dot{x},v) > 0$ for all $(t,x,u,\dot{x})$ in A and (t,x,u,v) in A.

b. $E(A) \subset C^1([t_0,t_1])$.

Proof: The assertion in (i) follows by Proposition 6, since $P(\cdot;x)$ is continuous on $[t_0,t_1]$. For (ii), let $G(t,v) := L_{\dot{x}}(t,x_0(t),x_0(t-\theta),v) - P(t;x_0) - c$, $g(t) := L_x(\tilde{x}_0(t))$ and $h(t) := L_u(\tilde{x}_0(t))$, where c is the constant obtained from Proposition 6. The assumptions imply that $G(t,\dot{x}_0(t)) = 0$ and $|G_v(t,\dot{x}_0(t))| \neq 0$ and, since,

$$G_t(t,v) = \frac{d}{dt}[L_{\dot{x}}(t,x_0(t),x_0(t-\theta),v)] - g(t) - h(t+\theta) \qquad t_0 \leqslant t \leqslant t_1-\theta$$

$$= \frac{d}{dt}[L_{\dot{x}}(t,x_0(t),x_0(t-\theta),v)] - g(t) \qquad t_1-\theta \leqslant t \leqslant t_1$$

$$G_v(t,v) = L_{\dot{x}\dot{x}}(t,x_0(t),x_0(t-\theta),v) \qquad t_0 \leqslant t \leqslant t_1,$$

the implicit function theorem clearly implies (a) and (b).

To prove (iii), observe first that, by Taylor's formula, given $(t,x,u,\dot{x})$ in A and (t,x,u,v) in A, there exists λ in (0,1) such that

$$E(t,x,u,\dot{x},v) = \frac{1}{2} \langle v - \dot{x} , L_{\dot{x}\dot{x}}(t,x,u,\dot{x} + \lambda(v - \dot{x}))(v - \dot{x})\rangle > 0$$

and (a) holds. For (b), let x in E(A) and suppose there exists s in (t_0,t_1) such that

$$a := \dot{x}(s - 0) \neq \dot{x}(s + 0) =: b.$$

Set, for all (t,v) in $[t_0,t_1] \times R^n$,

$$F(t,v) := L(t,x(t),x(t-\theta),v) - \langle v \ , \ L_{\dot{x}}(\tilde{x}(t))\rangle.$$

Since $L_{\dot{x}} \circ \tilde{x}$ is continuous on $[t_0,t_1]$,

$$F(s,a) - F(s,b) = E(s,x(s),x(s-\theta),b,a) > 0$$

$$F(s,b) - F(s,a) = E(s,x(s),x(s-\theta),a,b) > 0$$

and we arrive to a contradiction. Hence, $\dot{x}$ is continuous on $[t_0,t_1]$ and the proof is complete.

8. Corollary: Suppose L and $L_{\dot{x}\dot{x}}$ are continuous on A, x_0 is a non-singular trajectory and, for some $\varepsilon > 0$, $x_0 \in W(A;\varepsilon)$. Then ε can be diminished so that $E(t,x,u,\dot{x},v) > 0$ for all $(t,x,u,\dot{x})$ in $T_1(x_0;\varepsilon)$ and (t,x,u,v) in A with $v \neq \dot{x}$. Moreover, if x_0 is also an extremal, then x_0 is C^1. In other words, for all $\varepsilon > 0$, $E'(A) \cap W(A;\varepsilon) \subset C^1([t_0,t_1])$ and, in particular, $E'(A) \cap L(A) \subset C^1([t_0,t_1])$.

Proof: The first part follows by Lemma 4. It can also be proved directly as follows: diminish ε so that $|L_{\dot{x}\dot{x}}(t,x,u,\dot{x})| \neq 0$ for all $(t,x,u,\dot{x})$ in $T_1(x_0;\varepsilon)$. Suppose the result is false, i.e., there exist $\delta < \varepsilon$ and $(t,x,u,\dot{x},v)$ with $(t,x,u,\dot{x})$ in $T_1(x_0;\delta)$, (t,x,u,v) in A and $v \neq \dot{x}$, such that $E(t,x,u,\dot{x},u) = 0$. Setting $f(w) := E(t,x,u,w,v)$ for all w in R^n, f has a local minimum at $w = \dot{x}$ and so,

$$0 = f'(\dot{x}) = - L_{\dot{x}\dot{x}}(t,x,u,\dot{x})(v - \dot{x}).$$

But this implies $v = \dot{x}$ and the first part is proved.

Assume now that x_0 is also an extremal. As in the proof of Corollary 7(iii), suppose there exists s in (t_0,t_1) such that

$$a := \dot{x}_0(s - 0) \neq \dot{x}_0(s + 0) =: b.$$

Set, for all (t,v) in $[t_0,t_1] \times R^n$,

$$F(t,v) := L(t,x_0(t),x_0(t-\theta),v) - \langle v , L_{\dot{x}}(\tilde{x}_0(t))\rangle.$$

Since $L_{\dot{x}} \circ \tilde{x}_0$ is continuous on $[t_0,t_1]$ and, by definition, $(t,x_0(t),a)$ and $(t,x_0(t),b)$ belong to $T_1(x_0;\varepsilon) \cap A$,

$$F(s,a) - F(s,b) = E(s,x_0(s),x_0(s-\theta),b,a) > 0$$

$$F(s,b) - F(s,a) = E(s,x_0(s),x_0(s-\theta),a,b) > 0$$

which implies that x_0 is of class C^1.

Finally, assume only that $x_0 \in E'(A) \cap L(A)$. Since $L_{\dot{x}\dot{x}}$ is positive definite along x_0, there exists $\varepsilon > 0$ such that $|L_{\dot{x}\dot{x}}(t,x,u,\dot{x})| > 0$ for all $(t,x,u,\dot{x})$ in $T_1(x_0;\varepsilon)$. Let $(t,x,u,\dot{x})$ and (t,x,u,v) in $T_1(x_0;\varepsilon)$, and assume t in (t_0,t_1) does not correspond to a corner point of x_0. Since the set

$$\{(t,x,u,y,v) \mid |y - \dot{x}_0(t)| < \varepsilon\}$$

is convex, we obtain, by Taylor's formula,

$$E(t,x,u,\dot{x},v) = L(t,x,u,v) - L(t,x,u,\dot{x}) - \langle v - \dot{x} , L_{\dot{x}}(t,x,u,\dot{x})\rangle$$

$$= \int_0^1 (1 - \lambda)\{\langle v - \dot{x} , L_{\dot{x}\dot{x}}(t,x,u,\dot{x} + \lambda(v - \dot{x}))(v - \dot{x})\rangle\}d\lambda$$

$$\geqslant 0.$$

Therefore the condition of Weierstrass on $T_1(x_0;\varepsilon)$ holds along x_0 except possibly at the endpoints and corner points of x_0. But by continuity, it holds at these points also, and so $x_0 \in W(T_1(x_0;\varepsilon);\varepsilon)$. The previous result applies for this case and we obtain the required smoothness of x_0.

We prove next that nonsingular extremals satisfy a system of differential difference equations involving both advanced and retarded arguments.

9. Proposition: Suppose $L_{\dot{x}}$ is of class $C^r(A)$ $(r \geq 1)$, and x_0 is a nonsingular trajectory in $X(A)$. Set $p_0(t) := L_{\dot{x}}(\tilde{x}_0(t))$ and, for all $\varepsilon > 0$, let

$$T_2(x_0;\varepsilon) := \{(t,x,u,p) \in T_0(x_0;\varepsilon) \times R^n \mid |p - p_0(t)| < \varepsilon\}.$$

Then there exist $\varepsilon > 0$ and a function $\Lambda(t,x,u,p)$, with Λ_p of class C^r on $T_2(x_0;\varepsilon)$, such that the following are equivalent:

i. $x \in E'(A) \cap C^1([t_0,t_1])$, $p(t) = L_{\dot{x}}(\tilde{x}(t))$ and, for all t in $[t_0,t_1]$, $(t,x(t),x(t-\theta),p(t))$ belongs to $T_2(x_0;\varepsilon)$.

ii. (x,p) satisfies the difference-differential system (which we label (EE)):

$$\dot{x}(t) = \Lambda_p(t,x(t),x(t-\theta),p(t)) \qquad t_0 \leq t \leq t_1$$

$$\dot{p}(t) = -\Lambda_x(t,x(t),x(t-\theta),p(t)) - \Lambda_u(t+\theta,x(t+\theta),x(t),p(t+\theta)) \qquad t_0 \leq t \leq t_1-\theta$$

$$= -\Lambda_x(t,x(t),x(t-\theta),p(t)) \qquad t_1-\theta \leq t \leq t_1$$

Proof: By the implicit function theorem, there exist $\varepsilon > 0$ and a unique function U of class C^r mapping $T_2(x_0;\varepsilon)$ to R^n, such that, for all (t,x,u,p) in $T_2(x_0;\varepsilon)$,

$$p = L_{\dot{x}}(t,x,u,U(t,x,u,p)).$$

Define, for all (t,x,u,p) in $T_2(x_0;\varepsilon)$:

$$\Lambda(t,x,u,p) := \langle p\ ,\ U(t,x,u,p)\rangle - L(t,x,u,U(t,x,u,p))$$

(the **Legendre transform with respect to x_0**), and observe that

$$\Lambda_x(t,x,u,p) = - L_x(t,x,u,U(t,x,u,p))$$

$$\Lambda_u(t,x,u,p) = - L_u(t,x,u,U(t,x,u,p))$$

$$\Lambda_p(t,x,u,p) = U(t,x,u,p).$$

This implies the required smoothness of Λ. Now, clearly, (i) ⇒ (ii) follows by uniqueness of U and Proposition 6 and (ii) ⇒ (i) follows diminishing ε so that $|L_{\dot{x}\dot{x}}(t,x,u,\dot{x})| \neq 0$ for all $(t,x,u,\dot{x})$ in $T_1(x_0;\varepsilon)$.

5 THE SECOND VARIATION: JACOBI'S CONDITION

At this stage we can express, in terms of Euler's equation, trajectories for which the first variation vanishes on Y. Now we want to characterize H(A) and H′(A), i.e., the sets of trajectories for which the second variation is nonnegative or strictly positive in Y. In the delay free problem, they were characterized in terms of conjugate points, that is, points where the fundamental matrix solution of the Jacobi equation (Euler's equation for the integrand Ω), is singular. The whole purpose of this section is to find an equivalent concept for the delay problem.

Denote by J_x the second variation with respect to x and by E_x the set of secondary extremals with respect to x, i.e., trajectories satisfying Euler's equation for Ω:

$$J_x(y) := I''(x;y) = \int_{t_0}^{t_1} 2\Omega(t,y(t),y(t-\theta),\dot{y}(t))dt$$

$$E_x := \{y \in X \mid J_x'(y;z) = 0 \quad \forall z \in Y\}.$$

Several properties of these sets which will be used repeatedly throughout this section are summarized in the following lemma. They follow directly from the definitions.

10. Lemma: Suppose $L \in C^2(A;x,u,\dot{x})$. Then for all x, y and z in X, the following holds:

i. $2J_x(y) = J_x'(y;y)$

ii. $J_x'(y;z) = J_x'(z;y)$

iii. If $x \in H(A)$ and $y \in Y'$ then $y \in E_x$ if, and only if, $J_x(y) = 0$

iv. If $x \in H'(A)$ then $Y \cap E_x = \{0\}$.

Now, fix x_0 in X and set, for all $t_0 \leqslant t \leqslant t_1$,

$$A(t) := L_{\dot{x}\dot{x}}(\tilde{x}_0(t)) \qquad B(t) := L_{x\dot{x}}(\tilde{x}_0(t))$$
$$C(t) := L_{\dot{x}u}(\tilde{x}_0(t)) \qquad D(t) := L_{xx}(\tilde{x}_0(t))$$
$$E(t) := L_{xu}(\tilde{x}_0(t)) \qquad F(t) := L_{uu}(\tilde{x}_0(t))$$

so that, for all $(t,y,v,\dot{y})$ in $[t_0,t_1] \times R^{3n}$,

$$2\Omega(t,y,v,\dot{y}) := \langle \dot{y} , A(t)\dot{y}\rangle + \langle y , D(t)y\rangle + \langle v , F(t)v\rangle$$
$$+ 2\{\langle y , E(t)v\rangle + \langle y , B(t)\dot{y}\rangle + \langle \dot{y} , C(t)v\rangle\}, \qquad (2.18)$$

$$\Omega_y(t,y,v,\dot{y}) = B(t)\dot{y} + D(t)y + E(t)v$$
$$\Omega_v(t,y,v,\dot{y}) = C^*(t)\dot{y} + E^*(t)y + F(t)v$$
$$\Omega_{\dot{y}}(t,y,v,\dot{y}) = A(t)\dot{y} + B^*(t)y + C(t)v.$$

and

$$2\Omega(t,y,v,\dot{y}) = \langle y , \Omega_y(t,y,v,\dot{y})\rangle + \langle v , \Omega_v(t,y,v,\dot{y})\rangle + \langle \dot{y} , \Omega_{\dot{y}}(t,y,v,\dot{y})\rangle.$$

Assume that $x_0 \in L'(A)$. By Corollary 7(iii), all secondary extremals with respect to x_0 are of class C^1 and so $y \in E_{x_0}$ if, and only if,

$$\frac{d}{dt}\,\Omega_{\dot{y}}(\tilde{y}(t)) = \Omega_{y}(\tilde{y}(t)) + \Omega_{v}(\tilde{y}(t+\theta)) \qquad t_0 \leqslant t \leqslant t_1-\theta$$

$$= \Omega_{y}(\tilde{y}(t)) \qquad t_1-\theta \leqslant t \leqslant t_1.$$

This is Jacobi's equation for the trajectory x_0, which we label (JE). In terms of (2.18) it is given by:

$$\frac{d}{dt}[A(t)\dot{y}(t) + B^{*}(t)y(t) + C(t)y(t-\theta)]$$
$$= D(t)y(t) + E(t)y(t-\theta) + B(t)\dot{y}(t) + F(t+\theta)y(t)$$
$$+ E^{*}(t+\theta)y(t+\theta) + C^{*}(t+\theta)\dot{y}(t+\theta) \qquad t_0 \leqslant t \leqslant t_1-\theta$$
$$= D(t)y(t) + E(t)y(t-\theta) + B(t)\dot{y}(t) \qquad t_1-\theta < t \leqslant t_1.$$

Not surprisingly, Jacobi's equation for the delay problem presents shortcomings absent in the previous linear ordinary differential equation: the Legendre transform for the secondary problem is now given by a linear delay differential equation involving both advanced and retarded arguments. We start expressing (JE) in a standard form and see if we can apply general theorems on existence and uniqueness of solutions of the equation. Assuming A, B and C are differentiable, define $a_{ij}(t)$ for $i,j = 0,1,2$ as follows:

$$a_{00}(t) = \begin{cases} E^{*}(t) & t_0+\theta \leqslant t \leqslant t_1 \\ 0 & t_1 \leqslant t \leqslant t_1+\theta \end{cases}$$

$$a_{01}(t) = \begin{cases} C^{*}(t) & t_0+\theta \leqslant t \leqslant t_1 \\ 0 & t_1 \leqslant t \leqslant t_1+\theta \end{cases}$$

$$a_{10}(t) = \begin{cases} F(t) + D(t-\theta) - \dot{B}^{*}(t-\theta) & t_0+\theta \leqslant t \leqslant t_1 \\ D(t-\theta) - \dot{B}^{*}(t-\theta) & t_1 \leqslant t \leqslant t_1+\theta \end{cases}$$

and, for all t in $[t_0+\theta, t_1+\theta]$,

$$a_{20}(t) = E(t-\theta) - \dot{C}(t-\theta) \qquad a_{02}(t) \equiv a_{22}(t) \equiv 0$$

$$a_{11}(t) = B(t-\theta) - B^*(t-\theta) - \dot{A}(t-\theta)$$

$$a_{21}(t) = -C(t-\theta) \qquad a_{12}(t) = -A(t-\theta).$$

Setting $\omega_0 = 0$, $\omega_1 = \theta$ and $\omega_2 = 2\theta$, (JE) becomes, for all t in $[t_0+\theta, t_1+\theta]$,

$$\sum_{i=0}^{2} \sum_{j=0}^{2} a_{ij}(t) y^{(j)}(t-\omega_i) = 0. \tag{2.19}$$

Define now $A_i(t)$ and $B_i(t)$ $(i = 0,1,2)$ by:

$$A_0(t) = \begin{pmatrix} I & 0 \\ 0 & a_{02}(t) \end{pmatrix} \qquad B_0(t) = \begin{pmatrix} 0 & -I \\ a_{00}(t) & a_{01}(t) \end{pmatrix}$$

$$A_1(t) = \begin{pmatrix} 0 & 0 \\ 0 & a_{12}(t) \end{pmatrix} \qquad B_1(t) = \begin{pmatrix} 0 & 0 \\ a_{10}(t) & a_{11}(t) \end{pmatrix}$$

$$A_2(t) = \begin{pmatrix} 0 & 0 \\ 0 & a_{22}(t) \end{pmatrix} \qquad B_2(t) = \begin{pmatrix} 0 & 0 \\ a_{20}(t) & a_{21}(t) \end{pmatrix}$$

and observe that, if we set $v(t) = \begin{pmatrix} y(t) \\ \dot{y}(t) \end{pmatrix}$ then (2.19) is equivalent to the linear delay differential system:

$$\sum_{i=0}^{2} [A_i(t)\dot{v}(t-\omega_i) + B_i(t)v(t-\omega_i)] = 0 \qquad t_0+\theta \leqslant t \leqslant t_1+\theta. \tag{2.20}$$

For a system of this kind there do exist conditions implying existence and uniqueness of solutions. They are of the form:

$$v(t) = \psi(t) \qquad t_0-\theta \leqslant t \leqslant t_0+\theta \tag{2.21}$$

where $\psi(t)$ is a continuous function on $[t_0-\theta, t_0+\theta]$, the matrices $A_i(t)$ and $B_i(t)$ are continuous on $[t_0+\theta, t_1+\theta]$ and $A_0(t)$ is nonsingular on the entire interval $[t_0+\theta, t_1+\theta]$. Under these hypotheses, there exists a unique continuous function v(t) satisfying (2.20) and (2.21). Now, since $a_{02}(t) \equiv 0$, we have for our case that $\det(A_0(t)) = 0$. Several examples in [2] and [12] are given where, for this case and imposing initial conditions of the type (2.21), there does not exist a solution and, even if the matrices $A_i(t)$ and $B_i(t)$ are continuous and there exists a solution of (2.20) and (2.21), it may not necessarily be continuous.

The question at this point is whether there exist initial conditions for (JE) implying a unique continuous solution and, if any, see if they are appropriate for an equivalent definition of conjugate points. As we did in Chapter 1, let us express (JE) in terms of the Legendre transform for Ω. Observe that the assumption of A, B and C being differentiable is unnecessary in this representation.

11. Proposition: Suppose x_0 satisfies Legendre's strengthened condition. Then the following are equivalent:

i. y is a secondary extremal (with respect to x_0) and, for all t in $[t_0,t_1]$, $q(t) = A(t)\dot{y}(t) + B^*(t)y(t) + C(t)y(t-\theta)$.

ii. (y,q) satisfies the linear difference-differential system (which we label (JE)′):

1. $\dot{y}(t) = A_1(t)y(t) + A_2(t)y(t-\theta) + A_3(t)q(t)$

2a. $\dot{q}(t) = A_4(t)q(t) + A_5(t)q(t+\theta) + A_6(t)y(t) + A_7(t)y(t-\theta) + A_8(t)y(t+\theta)$

2b. $\dot{q}(t) = A_4(t)q(t) + A_9(t)y(t) + A_7(t)y(t-\theta)$.

(1) is satisfied on $[t_0,t_1]$, (2a) on $[t_0,t_1-\theta]$, (2b) on $[t_1-\theta,t_1]$ and the matrices A_i are given by:

$A_1(t) := -A^{-1}(t)B^*(t)$

$A_2(t) := -A^{-1}(t)C(t)$

$A_3(t) := A^{-1}(t)$

$A_4(t) := B(t)A^{-1}(t)$

$A_5(t) := C^*(t+\theta)A^{-1}(t+\theta)$

$A_6(t) := D(t) + F(t+\theta) - B(t)A^{-1}(t)B^*(t) - C^*(t+\theta)A^{-1}(t+\theta)C(t+\theta)$

$A_7(t) := E(t) - B(t)A^{-1}(t)C(t)$

$A_8(t) := E^*(t+\theta) - C^*(t+\theta)A^{-1}(t+\theta)B^*(t+\theta)$

$A_9(t) := D(t) - B(t)A^{-1}(t)B^*(t)$.

This is the equivalent of Proposition 1.9. As we shall see below, the difficulties mentioned above for finding a unique solution of (JE) are overcome imposing initial conditions for (JE)′ of the form

$$(1)' \quad y(t) = \psi(t) \qquad t_0-\theta \leqslant t \leqslant t_0$$

$$(2)' \quad q(c) = q_0$$

with $\psi:[t_0-\theta,t_0] \to \mathbf{R}^n$ continuous, q_0 in $\mathbf{R}^n$ and c any point in $[t_1-\theta,t_1]$.

12. **Lemma:** For all $h \geqslant 0$ let $\tau(h) := \max\{t_0, t_0 - \theta + h\}$, $\Omega_1(h) := C^0([t_0-\theta,t_0+h],\mathbf{R}^n)$, $\Omega_2(h) := C^0([t_0,t_0+h],\mathbf{R}^n)$ and suppose the matrices $A_1,\ldots,A_9$ are continuous. Then there exists h strictly positive such that, given $\psi:[t_0-\theta,t_0] \to \mathbf{R}^n$ continuous, $q_0 \in \mathbf{R}^n$ and $c \in [\tau(h),t_0+h]$, there is a unique $(y,q) \in \Omega_1(h) \times \Omega_2(h)$ solution of (1) on $[t_0,t_0+h]$, (2a) on $[t_0,\tau(h)]$ and (2b) on $[\tau(h),t_0+h]$, satisfying the initial conditions (1)′(2)′.

Proof: For all $i = 1,\ldots,9$ let $a_i := \| A_i \|$ and assume, without loss of generality, that there exists i such that $a_i \neq 0$. Let $M := (\sum_1^9 a_i)^{-1}$, take $0 < h < M$ with $h \leqslant t_1 - t_0$ and suppose $\psi:[t_0-\theta,t_0] \to \mathbf{R}^n$ continuous, $q_0 \in \mathbf{R}^n$ and $c \in [\tau(h),t_0+h]$ are given. We are going to prove that this way of choosing h satisfies the assertion of the lemma.

Let $S_1:\Omega_2(h) \to \Omega_1(h)$ and $S_2:\Omega_1(h) \to \Omega_2(h)$ be such that, for all

$q \in \Omega_2(h)$, $S_1(q)$ is the unique solution of (1) on $[t_0,t_0+h]$ satisfying (1)′ and, for all $y \in \Omega_2(h)$, $S_2(y)$ is the unique solution of (2a) on $[t_0,\tau(h)]$ and (2b) on $[\tau(h),t_0+h]$ satisfying (2)′. That these functions are well defined is easily verified in view of the theorem on existence and uniqueness of solutions for delay differential systems of this kind.

Now, if we set $S:= S_1 \circ S_2$, clearly the existence of a fixed point of S is equivalent to the existence of a solution of (1)(2) satisfying (1)′(2)′. Consequently, the lemma will be proved if we show that there exists $0 \leqslant \alpha < 1$ such that, for all $x,\ y \in \Omega_1(h)$, $\| S(x) - S(y)\| \leqslant \alpha\| x - y\|$ for, by the contraction principle, S will have a unique fixed point on $\Omega_1(h)$.

To begin with, note that the following holds for all y in $\Omega_1(h)$ and q in $\Omega_2(h)$:

$$(S_1(q))(t) = \psi(t) \qquad t_0-\theta \leqslant t \leqslant t_0 \tag{2.22}$$

$$= \psi(t_0) + \int_{t_0}^{t} [A_1(s)(S_1(q))(s) + A_2(s)(S_1(q))(s-\theta) + A_3(s)q(s)]ds \qquad t_0 \leqslant t \leqslant t_0+h$$

$$(S_2(y))(t) = \gamma(y;t) \qquad \tau(h) \leqslant t \leqslant t_0+h \tag{2.23}$$

$$= \gamma(y;\tau(h)) + \int_{\tau(h)}^{t} [A_4(s)(S_2(y))(s) + A_5(s)(S_2(y))(s+\theta) + A_6(s)y(s) + A_7(s)y(s-\theta) + A_8(s)y(s+\theta)]ds \qquad t_0 \leqslant t \leqslant \tau(h)$$

where, for all $t \in [\tau(h),t_0+h]$,

$$\gamma(y;t):= q_0 + \int_{c}^{t} [A_4(s)(S_2(y))(s) + A_9(s)y(s) + A_7(s)y(s-\theta)]ds.$$

Let $x,\ y \in \Omega_1(h)$. In view of (2.22),

$$\| S(x) - S(y)\| = \sup\{ \mid S_1(S_2(x))(t) - S_1(S_2(y))(t) \mid \mid t_0-\theta \leqslant t \leqslant t_0+h \}$$

$$\leqslant h(a_1+a_2)\| S(x) - S(y)\| + ha_3\| S_2(x) - S_2(y)\| .$$

This implies that, if $h(a_1+a_2) < 1$ and we set $\alpha_1 := ha_3/(1-h(a_1+a_2))$, then $\| S(x) - S(y)\| \leqslant \alpha_1\| S_2(x) - S_2(y)\|$.

On the other hand we have, in view of (2.23):

$$\| S_2(x) - S_2(y)\| = \sup\{ \mid (S_2(x))(t) - (S_2(y))(t) \mid \mid t_0 \leqslant t \leqslant t_0+h \}$$

$$\leqslant (\tau(h)-t_0)(a_4+a_5)\| S_2(x) - S_2(y)\| + (\tau(h)-t_0)(a_6+a_7+a_8)\| x - y\|$$

$$+ (t_0+h-\tau(h))(a_4\| S_2(x) - S_2(y)\| + (a_9+a_7)\| x - y\|)$$

$$= (ha_4 + (\tau(h)-t_0)a_5)\| S_2(x) - S_2(y)\|$$

$$+ [ha_7 + (t_0+h-\tau(h))a_9 + (\tau(h)-t_0)(a_6+a_8)]\| x - y\| .$$

So, if $ha_4 + (\tau(h) - t_0)a_5 < 1$, we have $\| S_2(x) - S_2(y)\| \leqslant \alpha_2\| x - y\|$, where:

$$\alpha_2 := \{h(a_7+a_9) + (\tau(h)-t_0)(a_6+a_8-a_9)\}/\{1 - (ha_4 + (\tau(h)-t_0)a_5)\}.$$

The proof is almost complete since, if h satisfies $h(a_1+a_2) < 1$ and $ha_4 + (\tau(h)-t_0)a_5 < 1$, then $\| S(x) - S(y)\| \leqslant \alpha\| x - y\|$, where $\alpha = \alpha_1\alpha_2$. It remains to show that the way we chose h implies these two conditions together with $0 \leqslant \alpha < 1$.

To this end observe that, since $h < (\sum_1^9 a_i)^{-1}$, $h(a_1 + a_2) < 1$ and the first condition holds. For the second, we may assume $\tau(h) = t_0-\theta+h > t_0$ since the case $\tau(h) = t_0$ is trivial. Hence, $\tau(h)-t_0 = h-\theta > 0$ which implies

$$ha_4 + (\tau(h)-t_0)a_5 = h(a_4+a_5) - \theta a_5 < 1$$

and the second condition follows. To prove that $0 \leqslant \alpha < 1$, observe first that $\alpha_1 < 1$ since $h(a_1+a_2+a_3) < 1$. Now, if $\tau(h) = t_0$, we clearly have

$\alpha_2 < 1$ and the result follows. Suppose then that $\tau(h) = t_0-\theta+h$. This implies:

$$h(a_7+a_9) + (\tau(h)-t_0)(a_6+a_8-a_9) + ha_4 + (\tau(h)-t_0)a_5$$

$$= h(a_4+a_7+a_9) + (\tau(h)-t_0)(a_5+a_6+a_8-a_9)$$

$$= h(a_4+a_5+a_6+a_7+a_8) - \theta(a_5+a_6+a_8-a_9)$$

$$< h(\sum_4^9 a_i) - \theta(a_5+a_6+a_8) < 1.$$

Hence, for this case we also obtain $\alpha_2 < 1$ and, by definition of α, it follows that $0 \leqslant \alpha < 1$. The proof is complete.

We are interested in obtaining Lemma 12 with the constant h replaced by $t_1 - t_0$, so that the unique solution of the Jacobi equation will be defined on the whole interval $[t_0-\theta,t_1]$. The proof of Lemma 12 implies the following result.

13. **Theorem:** Suppose $L \in C^2(A;x,u,\dot{x})$, $x_0 \in L'(A)$ is a trajectory of class C^1 on $[t_0,t_1]$ and

$$\theta \leqslant t_1 - t_0 < (\sum_1^9 \|A_i\|)^{-1}. \tag{2.24}$$

Then, for all $\psi:[t_0-\theta,t_0] \to R^n$ continuous, $q_0 \in R^n$ and $c \in [t_1-\theta,t_1]$, there exists a unique $(y,q) \in C^0([t_0-\theta,t_1],R^n) \times C^0([t_0,t_1],R^n)$ solution of (1)(2), satisfying the initial data (1)′(2)′.

Remark: Throughout the remaining of this chapter we shall assume that condition (2.24) holds, thus limiting ourselves to problems for which the delay is "sufficiently small". It should be noted that, under the usual smoothness hypotheses for the Legendre transform, an analogous proof to that of Lemma 12 shows that initial conditions (1)′(2)′

guarantee existence and uniqueness of solutions of (EE).

Based on Theorem 13, the notion of "conjugate point" seems now natural: denote by $(Y(\cdot,t_0),Q(\cdot,t_0))$ the (matrix) solution of (JE)′ satisfying:

$$Y(t,t_0) = 0 \qquad \text{for} \quad t_0-\theta \leqslant t \leqslant t_0$$

$$Q(t_1,t_0) = I.$$

Following the definition for the problem without delays, a point s in $(t_0,t_1]$ would be called conjugate to t_0 with respect to x_0 if $Y(s,t_0)$ is degenerate. As the following proposition shows, there would be a slight difference between the two definitions.

14. **Lemma:** Let s be any point in $(t_0,t_1]$. Then the following statements are equivalent:

i. $Y(s,t_0)$ is singular.

ii. There exists y, a nonvanishing secondary extremal with respect to x_0, satisfying $y(t) = 0$ for all t in $[t_0-\theta,t_0]$ and $y(s) = 0$.

Proof: (i) $\Rightarrow$ (ii) follows clearly defining $y(t) := Y(t,t_0)c$ where c in $R^n - \{0\}$ is such that $Y(s,t_0)c = 0$. Hence, assume (ii) is true and let, for all t in $[t_0,t_1]$,

$$q(t) := A(t)\dot{y}(t) + B^*(t)y(t) + C(t)y(t-\theta).$$

Let $c := q(t_1)$. Since $y(t) = 0$ on $[t_0-\theta,t_0]$, $c \neq 0$ for otherwise $y \equiv 0$. Define $z(t) := Y(t,t_0)c$ and $r(t) := Q(t,t_0)c$. Hence, (z,r) satisfies (JE)′, z vanishes on $[t_0-\theta,t_0]$ and $r(t_1) = c = q(t_1)$, which implies that

$(z,r) = (y,q)$. Consequently, $Y(s,t_0)c = y(s) = 0$ and (i) follows.

The slight difference is that it cannot be assured the nonvanishing of the secondary extremal y on the whole interval (t_0,s). As we saw in the proof of Proposition 1.10, this played a fundamental role in order to characterize H(A) in terms of conjugate points. The next lemma shows that there is an even stronger difference between the two problems.

15. Lemma: Suppose there exist y in E_{x_0} vanishing on $[t_0-\theta,t_0]$ and c in $(t_0,t_1]$ such that $y(c) = 0$. Let

$$z(t):= \begin{cases} y(t) & t_0-\theta \leqslant t \leqslant c \\ 0 & c \leqslant t \leqslant t_1 \end{cases}$$

Then z belongs to Y and

$$J_{x_0}(z) = -\int_c^{\tau(c)} \langle y(t-\theta) \ , \ C^*(t)\dot{y}(t) + E^*(t)y(t)\rangle dt$$

where $\tau(t):= \min\{t+\theta,t_1\}$.

Proof: From the definition of Ω it follows that:

$$\begin{aligned} \Omega_y(\tilde{z}(t)) &= \Omega_y(\tilde{y}(t)) && t_0 \leqslant t < c \\ &= E(t)y(t-\theta) && c < t \leqslant \tau(c) \\ &= 0 && \tau(c) \leqslant t \leqslant t_1 \end{aligned}$$

$$\begin{aligned} \Omega_{\dot{y}}(\tilde{z}(t)) &= \Omega_{\dot{y}}(\tilde{y}(t)) && t_0 \leqslant t < c \\ &= C(t)y(t-\theta) && c < t \leqslant \tau(c) \\ &= 0 && \tau(c) \leqslant t \leqslant t_1 \end{aligned}$$

$$\Omega_v(\tilde{z}(t)) = \Omega_v(\tilde{y}(t)) \qquad t_0 \leqslant t < c$$
$$= F(t)y(t-\theta) \qquad c < t \leqslant \tau(c)$$
$$= 0 \qquad \tau(c) \leqslant t \leqslant t_1$$

Consequently,

$$J_{x_0}(z) = \int_{t_0}^{t_1} \{\langle z(t) , \Omega_y(\tilde{z}(t))\rangle + \langle z(t-\theta) , \Omega_v(\tilde{z}(t))\rangle + \langle \dot{z}(t) , \Omega_{\dot{y}}(\tilde{z}(t))\rangle\}dt$$

$$= \int_{t_0}^{c} \{\langle y(t) , \Omega_y(\tilde{y}(t))\rangle + \langle y(t-\theta) , \Omega_v(\tilde{y}(t))\rangle + \langle \dot{y}(t) , \Omega_{\dot{y}}(\tilde{y}(t))\rangle\}dt$$

$$+ \int_{c}^{\tau(c)} \langle y(t-\theta) , F(t)y(t-\theta)\rangle dt$$

For the case $c \in (t_0, t_1-\theta]$ we obtain:

$$J_{x_0}(z) = \int_{t_0}^{c} \{\langle y(t) , \Omega_y(\tilde{y}(t)) + \Omega_v(\tilde{y}(t+\theta))\rangle + \langle \dot{y}(t) , \Omega_{\dot{y}}(\tilde{y}(t))\rangle\}dt$$

$$+ \int_{c}^{c+\theta} \langle y(t-\theta) , F(t)y(t-\theta)\rangle dt - \int_{c-\theta}^{c} \langle y(t) , \Omega_v(\tilde{y}(t+\theta))\rangle dt$$

$$= \int_{t_0}^{c} \frac{d}{dt}[\langle y(t) , \Omega_{\dot{y}}(\tilde{y}(t))\rangle]dt$$

$$+ \int_{c}^{c+\theta} \langle y(t-\theta) , F(t)y(t-\theta) - \Omega_v(\tilde{y}(t))\rangle dt$$

$$= - \int_{c}^{c+\theta} \langle y(t-\theta) , C^*(t)\dot{y}(t) + E^*(t)y(t)\rangle dt$$

For the case $c \in (t_1-\theta, t_1]$ the proof is similar and we obtain:

$$J_{x_0}(z) = - \int_{c}^{t_1} \langle y(t-\theta) , C^*(t)\dot{y}(t) + E^*(t)y(t)\rangle dt$$

In view of the last two lemmas, we adopt the following definition of conjugate point:

Definition: A point c in $(t_0,t_1]$ will be called **conjugate to t_0 with respect to x_0** if there exists y in E_{x_0} such that:

i. y is nonvanishing on (t_0,c).

ii. $y(t) = 0$ for all t in $[t_0-\theta,t_0]$ and $y(c) = 0$.

iii. $\int_c^{\tau(c)} \langle y(t-\theta)\ ,\ C^*(t)\dot{y}(t) + E^*(t)y(t)\rangle dt \geq 0.$

where $\tau(t) := \min\{t+\theta,t_1\}$.

According to this definition, we characterize the set H(A) in the next proposition.

16. **Proposition:** Suppose $L \in C^2(A;x,u,\dot{x})$ and x_0 is a C^1 trajectory in H(A) satisfying the strengthened Legendre condition. Then the following holds:

a. There are no conjugate points to t_0 on the interval $(t_0,t_1-\theta]$.

b. If there exists s in $(t_1-\theta,t_1]$ such that $L_{\dot{x}u}(x_0(s)) = 0$, then there are no conjugate points to t_0 on (t_0,s).

Proof: Suppose there exist $c \in (t_0,t_1-\theta]$ and y secondary extremal, satisfying (i)-(iii) above. In view of (ii) and (iii) and Lemma 15, the function

$$z(t) := \begin{cases} y(t) & t_0-\theta \leq t \leq c \\ 0 & c \leq t \leq t_1 \end{cases}$$

belongs to Y and $J_{x_0}(z) = 0$. Since $x_0 \in H(A)$, z must be a secondary extremal with respect to x_0 and so it satisfies (JE). Setting

$$r(t) := A(t)\dot{z}(t) + B^*(t)z(t) + C(t)z(t-\theta)$$

(z,r) satisfies (JE)'. Now, $z(t) = 0$ on $[t_0-\theta,t_0]$ and, since

$c \in (t_0, t_1-\theta]$, $r(t) = 0$ for all t in $[c+\theta, t_1]$. By Theorem 13, it follows that $z \equiv 0$, thus contradicting (i). This proves (a) and, clearly, the same argument applies for (b).

17. Remark: Consider the following two conditions:

C1. If y satisfies (JE) and there exists $c > t_0$ such that $y(t) = 0$ for all t in $[t_0-\theta, c]$, then $y \equiv 0$.

C2. The matrix

$$K_Y(t) := \int_t^{\tau(t)} Y^*(s-\theta, t_0)(E^*(s)Y(s,t_0) + C^*(s)\dot{Y}(s,t_0))ds$$

is positive semidefinite at points in $(t_0, t_1]$ where $Y(\cdot, t_0)$ is singular, i.e., $|Y(c,t_0)| = 0 \Rightarrow K_Y(c) \geqslant 0$.

Observe that, assuming these conditions, Proposition 16 holds if we define conjugate points as was suggested in the previous "natural" way, i.e., points where the matrix $Y(\cdot, t_0)$ is singular. In fact, under these conditions, the two definitions are equivalent in view of Lemma 14.

Conditions (a) and (b) of Proposition 1.16 will be called **Jacobi's condition.** We turn now to characterize condition H′(A) and find an analog of Jacobi's strengthened condition. For the delay free problem, it was shown in Proposition 1.12 that, if (Y,Q) is a solution of (JE)′, then $Y^*(t)Q(t) - Q^*(t)Y(t)$ is constant. This was basic in order to prove that $x_0 \in H'(A)$ if, and only if, the matrix $Y(\cdot, t_0)$ is nonsingular on the interval $(t_0, t_1]$. The next lemma shows what happens in the delay case.

18. Lemma: Suppose $L \in C^2(A; x, u, \dot{x})$, (Y,Q) is a matrix solution of (JE)′ and $Y(t) = 0$ for all t in $[t_0-\theta, t_0]$. Then:

$$Q^*(t)Y(t) - Y^*(t)Q(t) = \int_t^{\tau(t)} (\dot{Y}^*(s)C(s) + Y^*(s)E(s))Y(s-\theta)ds$$

$$- \int_t^{\tau(t)} Y^*(s-\theta)(C^*(s)\dot{Y}(s) + E^*(s)Y(s))ds \qquad (2.25)$$

where $\tau(t) = \min\{t+\theta, t_1\}$.

Proof: For all y and z in X let:

$$K(y,z;t) := \int_t^{\tau(t)} \langle z(s-\theta) , C^*(s)\dot{y}(s) + E^*(s)y(s)\rangle ds.$$

We shall first prove that, if y and z are any trajectories in E_{x_0} then, for all t in $[t_0,t_1]$:

$$\langle z(t) , \Omega_{\dot{y}}(\tilde{y}(t))\rangle - \langle z(t_0) , \Omega_{\dot{y}}(\tilde{y}(t_0))\rangle + K(y,z;t_0) - K(y,z;t)$$

$$= \langle y(t) , \Omega_{\dot{y}}(\tilde{z}(t))\rangle - \langle y(t_0) , \Omega_{\dot{y}}(\tilde{z}(t_0))\rangle$$

$$+ K(z,y;t_0) - K(z,y;t) \qquad (2.26)$$

Let t in $[t_0,t_1]$ and assume $t \leqslant t_1-\theta$ (for the case $t > t_1-\theta$, the proof is analogous). From the definitions we have:

$$\langle z(t) , \Omega_{\dot{y}}(\tilde{y}(t))\rangle - \langle z(t_0) , \Omega_{\dot{y}}(\tilde{y}(t_0))\rangle$$

$$= \int_{t_0}^{t} \{\langle z(s) , \Omega_y(\tilde{y}(s)) + \Omega_v(\tilde{y}(s+\theta))\rangle + \langle \dot{z}(s) , \Omega_{\dot{y}}(\tilde{y}(s))\rangle\}ds$$

$$= \langle y(t) , \Omega_{\dot{y}}(\tilde{z}(t))\rangle - \langle y(t_0) , \Omega_{\dot{y}}(\tilde{z}(t_0))\rangle$$

$$+ \int_{t_0}^{t_0+\theta} \{\langle y(s-\theta) , \Omega_v(\tilde{z}(s))\rangle - \langle z(s-\theta) , \Omega_v(\tilde{y}(s))\rangle\}ds$$

$$+ \int_{t}^{t+\theta} \{\langle z(s-\theta) , \Omega_v(\tilde{y}(s))\rangle - \langle y(s-\theta) , \Omega_v(\tilde{z}(s))\rangle\}ds$$

which implies (2.26).

Now, denote the columns of Y and Q by y_i and q_i respectively. Since $q_i(t) = \mathcal{Q}_{\dot{y}}(\tilde{y}_i(t))$, the (i,j)-entry of the matrix $Q^*(t)Y(t) - Y^*(t)Q(t)$ is given by

$$\langle q_i(t) , y_j(t)\rangle - \langle y_i(t) , q_j(t)\rangle$$

$$= K(y_i,y_j;t) - K(y_j,y_i;t)$$

$$= \int_t^{\tau(t)} \{\langle \dot{y}_i(s) , C(s)y_j(s-\theta)\rangle + \langle y_i(s) , E(s)y_j(s-\theta)\rangle\}ds$$

$$- \int_t^{\tau(t)} \{\langle y_i(s-\theta) , C^*(s)\dot{y}_j(s) + E^*(s)y_j(s)\rangle\}ds$$

which is the (i,j)-entry of the right hand matrix in (2.25). The equality follows.

19. Lemma: Suppose there exists a matrix solution (Y,Q) of (JE)′ such that $|Y(t)| \neq 0$ and

$$Y^*(t)Q(t) - Q^*(t)Y(t) = \int_t^{\tau(t)} [R^*(s) - R(s)]ds$$

for all t in $[t_0,t_1]$, where $R(t) = Y^*(t-\theta)(C^*(t)\dot{Y}(t) + E^*(t)Y(t))$ and $\tau(t) = \min\{t+\theta,t_1\}$. For all y in Y, let

$$w(t):= \begin{cases} 0 & t_0-\theta \leqslant t \leqslant t_0 \\ Y^{-1}(t)y(t) & t_0 \leqslant t \leqslant t_1 \end{cases}$$

and

$$T(t):= Y^*(t-\theta)(C^*(t)\dot{y}(t) + E^*(t)y(t)).$$

Then

$$J_{x_0}(y) = \int_{t_0}^{t_1} \langle Y(t)\dot{w}(t) , L_{\dot{x}\dot{x}}(\tilde{x}_0(t))Y(t)\dot{w}(t)\rangle dt$$

$$+ 2\int_{t_0}^{t_1} \langle \dot{w}(t) , \int_{t}^{\tau(t)} \{R(s)w(t) - T(s)\}ds\rangle dt.$$

Proof: The proof follows directly from the definitions. Let $z(t) := Y(t)\dot{w}(t)$ so that $\dot{y}(t) = \dot{Y}(t)w(t) + z(t)$. From the definition of Ω,

$$\langle \dot{y}(t) , \Omega_{\dot{y}}(\tilde{y}(t))\rangle = \langle \dot{Y}(t)w(t) + z(t) , Q(t)w(t) + A(t)z(t)\rangle$$

$$+ \langle \dot{Y}(t)w(t) + z(t) , C(t)Y(t-\theta)(w(t-\theta) - w(t))\rangle$$

and similarly, for $t_0 \leqslant t \leqslant t_1-\theta$,

$$\langle y(t), \Omega_y(\tilde{y}(t)) + \Omega_v(\tilde{y}(t+\theta))\rangle = \langle Y(t)w(t) , \dot{Q}(t)w(t) + B(t)z(t) + C^*(t+\theta)z(t+\theta)\rangle$$

$$+ \langle w(t) , Y^*(t)E(t)Y(t-\theta)(w(t-\theta) - w(t)) + R(t+\theta)(w(t+\theta) - w(t))\rangle$$

and, for $t_1-\theta \leqslant t \leqslant t_1$,

$$\langle y(t) , \Omega_y(\tilde{y}(t))\rangle = \langle Y(t)w(t) , \dot{Q}(t)w(t) + B(t)z(t)\rangle$$

$$+ \langle Y(t)w(t) , E(t)Y(t-\theta)(w(t-\theta) - w(t))\rangle.$$

Hence,

$$J_{x_0}(y) = \int_{t_0}^{t_1} \{\frac{d}{dt}(\langle w(t) , Q^*(t)Y(t)w(t)\rangle) + \langle z(t) , A(t)z(t)\rangle\}dt$$

$$+ 2\int_{t_0}^{t_1} \langle w(t-\theta) - w(t) , Y^*(t-\theta)C^*(t)z(t) + R(t)w(t)\rangle dt$$

$$+ \int_{t_0}^{t_1} \langle w(t) , R(t)w(t) + (Q^*(t)Y(t) - Y^*(t)Q(t))\dot{w}(t)\rangle$$

$$- \int_{t_0}^{t_1} \langle w(t-\theta) , R(t)w(t-\theta)\rangle dt$$

$$= \int_{t_0}^{t_1} \langle z(t) , A(t)z(t)\rangle$$

$$+ 2 \int_{t_0}^{t_1} \{\langle w(t-\theta) - w(t) , T(t)\rangle + \langle \dot{w}(t) , (\int_{t}^{\tau(t)} R(s)ds)w(t)\rangle\}dt$$

and the result follows.

Consider the following condition:

C3. For all y in Y, $\int_{t_0}^{t_1} \langle \dot{w}(t) , \int_{t}^{\tau(t)} \{R(s)w(t) - T(s)\}ds\rangle dt \geq 0.$

We are now in a position to prove the analog of Proposition 1.12. We shall say that x_0 satisfies **Jacobi's strengthened condition** if there are no conjugate points to t_0 with respect to x_0 on $(t_0,t_1]$.

20. **Proposition:** Suppose $L \in C^2(A;x,u)$, $L_{\dot{x}} \in C^1(A)$ and x_0 is a nonsingular extremal satisfying the Legendre condition. If conditions (C1), (C2) and (C3) hold, then the following are equivalent:

i. $x_0 \in H'(A)$.

ii. x_0 satisfies Jacobi's strengthened condition.

iii. The matrix $Y(t,t_0)$ is nonsingular for all t in $(t_0,t_1]$.

Proof: (i) ⇒ (ii): By Corollary 8, x_0 is C^1 on $[t_0,t_1]$ and so, by Proposition 16, (ii) holds on the interval $(t_0,t_1-\theta]$. Suppose there exists a conjugate point on $(t_1-\theta,t_1]$. Then, as in the proof of Proposition 16, there exists z in Y - {0} such that $J_{x_0}(z) \leq 0$ contradicting (i).

(ii) ⇒ (iii): See Remark 17.

(iii) → (i): Setting $p_0(t) = L_{\dot{x}}(\tilde{x}_0(t))$, (x_0,p_0) satisfies (EE). In the proof of Proposition 9 we saw that solutions of (EE) are C^1-extremals lying in a neighborhood of $(t,x_0(t),x_0(t-\theta),p_0(t))$, and so (x_0,p_0) can be extended over a larger interval $[t_0-\varepsilon,t_1+\varepsilon]$. Solve the problem for (EE) (see the remark following Theorem 13) with initial data $(0 < \delta < \varepsilon)$:

$$x(t,\lambda) = x_0(t) \qquad t_0-\theta-\delta \leqslant t \leqslant t_0-\delta$$

$$p(t_1,\lambda) = \lambda + p_0(t_1).$$

Since $x(\cdot,\lambda)$ are extremals of L, the following holds:

$$-\frac{d}{dt}[L_{\dot{x}}(\tilde{x}(t,\lambda))] + L_x(\tilde{x}(t,\lambda)) + L_u(\tilde{x}(t+\theta,\lambda)) \equiv 0 \qquad t_0 \leqslant t \leqslant t_1-\theta$$

$$-\frac{d}{dt}[L_{\dot{x}}(\tilde{x}(t,\lambda))] + L_x(\tilde{x}(t,\lambda)) \equiv 0 \qquad t_1-\theta \leqslant t \leqslant t_1.$$

The smoothness of $L_{\dot{x}}$ permits to differentiate with respect to λ. Since $x(t,0) = x_0(t)$ by uniqueness, it follows that the matrix $Y(t,t_0-\delta)$: $\partial x(t,0)/\partial\lambda$ satisfies Jacobi equation (JE). Setting

$$Q(t,t_0-\delta) := A(t)\dot{Y}(t,t_0-\delta) + B^*(t)Y(t,t_0-\delta) + C(t)Y(t-\theta,t_0-\delta),$$

it follows that $(Y(t,t_0-\delta),Q(t,t_0-\delta))$ solves (JE)′.

Observe now that $Q(t,t_0-\delta) = \partial p(t,0)/\partial\lambda$ since, by Proposition 9, $p(t,\lambda) = L_{\dot{x}}(\tilde{x}(t,\lambda))$. By the boundary conditions imposed for $x(\cdot,\lambda)$ and $p(\cdot,\lambda)$,

$$Y(t,t_0-\delta) = 0 \qquad t_0-\theta-\delta \leqslant t \leqslant t_0-\delta$$

$$Q(t_1,t_0-\delta) = I.$$

In view of (iii) it follows, by continuity considerations, that the matrix $Y(t,t_0-\delta)$ is nonsingular for some $\delta > 0$ on the entire interval

$(t_0-\delta,t_1]$. The rest of the proof is the content of Lemmas 18 and 19.

We summarize the results of this and the previous section stating the analogs of Theorems 1.11 and 1.13. They follow directly from Propositions 6, 16 and 20.

21. **Theorem:** Suppose $L \in C^2(A;x,u,\dot{x})$ and x is a trajectory solving P(A). Then x satisfies Euler's equation and the conditions of Legendre and Weierstrass. If also x is nonsingular then x satisfies Jacobi's condition.

22. **Theorem:** Suppose $L \in C^2(A;x,u)$, $L_{\dot{x}} \in C^1(A)$ and conditions (C1), (C2) and (C3) hold. If x in $X_{\bullet}(A)$ is a nonsingular trajectory satisfying Euler's equation, Legendre's condition and Jacobi's strengthened condition, then x is a weak minimum for P(A). If also L is regular on A or x satisfies the strengthened condition of Weierstrass, then x is a strong minimum for P(A).

Remark: It should be noted that the proof of Proposition 20 permits us to embed explicitly trajectories belonging to H'(A) into families of extremals, suggesting an extension of Mayer fields and the Hamilton-Jacobi theory. The conditions required for these extensions turn out to be quite similar to those of (C3). Consequently, we shall stop at this point the parallel between the two problems. The shortcomings presented in order to verify these conditions will be overcome in the next chapter through a completely different approach.

CHAPTER 3

THE METHOD OF STEPS

1 INTRODUCTION

The problem studied in Chapter 2 is converted, through the "method of steps", into one without delays. This new problem will not have fixed endpoint constraints but necessary conditions for a minimum will turn out to be a direct consequence of the results obtained in Chapter 1. Sufficiency will be derived through the Hamilton-Jacobi inequality, fields of extremals and the positivity of the second variation, adding an extra condition to the classical sufficient conditions for delay free problems. This extra condition is based exclusively on a solution of a given matrix Riccati equation. The main difference between these conditions and those stated in Theorems 2.21 and 2.22, will be the concept of conjugate point. We introduce the notion of "conjugate sequence" and overcome the difficulties presented in order to verify the existence of the previous conjugate points.

We end this chapter presenting two examples illustrating some of the results. For the first one we use the notion of conjugate point to show that the problem under consideration has no minimum. For the second example, consisting of a problem with a weak minimum which is not a strong minimum, we apply the sufficient conditions of Theorem 2.5.

2 AN EQUIVALENT NON-DELAY PROBLEM

We shall be dealing with the problem P(A) of Chapter 2 with $A = [t_0,t_1] \times R^{3n}$. We assume throughout this chapter that the Lagrangian is integrable along any trajectory and the delay is strictly positive. For ease of notation every concept will be written without referring explicitly to the set A. Hence, our problem, labeled (P), will be that of minimizing the functional $I(x) = \int_{t_0}^{t_1} L(\tilde{x}(t))dt$ over $X_\bullet$, where the space of trajectories is given by

$$X = \{x:[t_0-\theta,t_1] \to R^n \mid x \text{ is piecewise-}C^1\},$$

and the endpoint constraints are

$$X_\bullet = \{x \in X \mid x(t) = \phi(t) \quad \forall\ t \in [t_0-\theta,t_0] \text{ and } x(t_1) = \xi\}.$$

In this section we convert (P) into a non-delay problem. Let us start defining formally in what sense two problems will be called equivalent.

1. Definition: Given two sets X and $\hat{X}$ and functionals $I:X \to R$ and $\hat{I}:\hat{X} \to R$, consider the problems:

$$\text{P:} \quad \text{Minimize } I(x) \text{ over } X_\bullet$$
$$\hat{\text{P}}\text{:} \quad \text{Minimize } \hat{I}(\hat{x}) \text{ over } \hat{X}_\bullet,$$

where $X_\bullet$ and $\hat{X}_\bullet$ are subsets of X and $\hat{X}$ respectively. We shall say (P) and ($\hat{P}$) are **equivalent** if there exists a one-to-one mapping Φ of $X_\bullet$ onto $\hat{X}_\bullet$ which leaves the value of the functionals unchanged, that is, one for which $I(x) = \hat{I}(\Phi(x))$ for all $x \in X_\bullet$.

Clearly, if two problems (P) and ($\hat{P}$) are equivalent, x solves (P) if, and only if, $\Phi(x)$ solves ($\hat{P}$) and, when X and $\hat{X}$ are normed spaces and Φ and Φ^{-1} are continuous with respect to the norm topology, the same applies for local optimality.

In order to find an equivalent non-delay problem for (P), consider the following definitions:

i. Let $N := \max\{k \in \mathbf{N} \mid t_0 + k\theta < t_1\}$, $p := N + 1$ and $t_2 := t_0 + p\theta$.

ii. Define $\hat{L}:[0,\theta] \times \mathbf{R}^{2np} \to \mathbf{R}$ as follows: extend L such that, for all t in $(t_1,t_2]$ and $(x,u,\dot{x})$ in $\mathbf{R}^{3n}$, $L(t,x,u,\dot{x}) = L(t_1,x,u,\dot{x})$ and let, for all t in $[0,\theta]$ and $\hat{x} = (\hat{x}_0,\ldots,\hat{x}_N)$ and $\hat{v} = (\hat{v}_0,\ldots,\hat{v}_N)$ in $\mathbf{R}^{np}$,

$$\hat{L}(t,\hat{x},\hat{v}) := L(t+t_0,\hat{x}_0,\phi(t+t_0-\theta),\hat{v}_0) + \sum_{k=1}^{N} L(t+t_0+k\theta,\hat{x}_k,\hat{x}_{k-1},\hat{v}_k).$$

iii. Define $c:\mathbf{R}^{np} \to \mathbf{R}^{np}$, $\hat{X}$ and $\hat{X}_\bullet$ as follows:

$$c(\hat{x}_0,\ldots,\hat{x}_N) = (\phi(t_0),\hat{x}_0,\ldots,\hat{x}_{N-1}) \quad \text{for all } (\hat{x}_0,\ldots,\hat{x}_N) \in \mathbf{R}^{np}$$

$$\hat{X} := \{\hat{x}:[0,\theta] \to \mathbf{R}^{np} \mid \hat{x} \text{ is piecewise-}C^1\}$$

$$\hat{X}_\bullet := \{\hat{x} \in \hat{X} \mid \hat{x}(0) = c(\hat{x}(\theta)) \text{ and } \hat{x}_N(t-t_0-N\theta) = \xi \quad \forall\, t \in [t_1,t_2]\}.$$

Consider now the new problem:

$$\hat{P}: \quad \text{Minimize } \hat{I}(\hat{x}) \text{ over } \hat{X}_\bullet$$

where the functional $\hat{I}:\hat{X} \to \mathbf{R}$ is given by

$$\hat{I}(\hat{x}) := \int_0^\theta \hat{L}(t,\hat{x}(t),\dot{\hat{x}}(t))dt \quad \text{for all } \hat{x} \in \hat{X}.$$

2. **Proposition:** Problems (P) and ($\hat{P}$) are equivalent.

Proof: Given x in X, extend x such that $x(t) = x(t_1)$ for all t in $(t_1,t_2]$ and let $\Phi = (\Phi_0,\ldots,\Phi_N):X \to \hat{X}$ be such that

$$(\Phi_k(x))(t) := x(t+t_0+k\theta) \qquad k = 0,1,\ldots,N, \qquad t \in [0,\theta].$$

By construction, it is clear that $\Phi(X_\bullet) = \hat{X}_\bullet$ and Φ (restricted to $X_\bullet$) is one-to-one. The inverse of Φ (restricted to $X_\bullet$) is explicitly given by:

$$
\begin{aligned}
(\Phi^{-1}(\hat{x}))(t) &= \phi(t) && t \in [t_0-\theta, t_0] \\
&= \hat{x}_k(t-t_0-k\theta) && t \in [t_0+k\theta, t_0+(k+1)\theta] \quad k = 0,1,\ldots,N-1 \\
&= \hat{x}_N(t-t_0-N\theta) && t \in [t_0+N\theta, t_1].
\end{aligned}
$$

To prove the invariance under the integral, let $x \in X_\bullet$ and set $\hat{x} := \Phi(x)$. From the definition of the functionals, we have:

$$
\begin{aligned}
I(x) &= \int_{t_0}^{t_1} L(\tilde{x}(t))dt \\
&= \int_0^\theta \sum_{k=0}^N L(\tilde{x}(t+t_0+k\theta))dt \\
&= \int_0^\theta L(t+t_0, \hat{x}_0(t), \phi(t+t_0-\theta), \dot{\hat{x}}_0(t))dt \\
&\qquad + \int_0^\theta \sum_{k=1}^N L(t+t_0+k\theta, \hat{x}_k(t), \hat{x}_{k-1}(t), \dot{\hat{x}}_k(t))dt \\
&= \hat{I}(\hat{x}).
\end{aligned}
$$

In a natural way we extend the concepts of tubes and restricted tubes, for the space $\hat{X}$, as those defined in Chapter 1. For example:

$$\hat{T}_0(\hat{x};\varepsilon) = \{(t,\hat{y}) \in [0,\theta] \times R^{np} \mid |\hat{x}(t) - \hat{y}| < \varepsilon\}.$$

By construction, it is clear that the function Φ preserves local optimality. Now, since (P) and ($\hat{P}$) are equivalent, we obtain immediately necessary conditions for (P) for, if $\hat{x}$ solves ($\hat{P}$), then $\hat{x}$ also solves the problem:

$$\hat{P}': \quad \text{Minimize } \hat{I}(\hat{y}) \text{ over } \hat{X}_\bullet'$$

where:

$$\hat{X}_\bullet' := \{\hat{y} \in \hat{X} \mid \hat{y}(0) = \hat{x}(0) \text{ and } \hat{y}(\theta) = \hat{x}(\theta)\}$$

and, for this fixed endpoint problem, we can apply the results of Chapter 1.

These remarks and Theorem 1.3 imply the following result. Denote by $\hat{Y}$ the set of trajectories in $\hat{X}$ vanishing at 0 and θ and the Weierstrass excess function for $\hat{L}$ by $\hat{E}$.

3. **Theorem:** Suppose $\hat{L} \in C^2(\hat{x},\hat{v})$ and x is a weak minimum for (P). Let $\hat{x} := \Phi(x)$. Then the following holds:

i. For all $\hat{y} \in \hat{Y}$, $\hat{I}'(\hat{x};\hat{y}) = 0$.

ii. For all $\hat{y} \in \hat{Y}$, $\hat{I}''(\hat{x};\hat{y}) \geq 0$.

iii. For all $t \in [0,\theta]$, $\hat{L}_{\hat{v}\hat{v}}(t,\hat{x}(t),\dot{\hat{x}}(t)) \geq 0$.

If x is a strong minimum for (P), then $\hat{x}$ also satisfies:

iv. For all $\hat{u} \in R^{np}$, $\hat{E}(t,\hat{x}(t),\dot{\hat{x}}(t),\hat{u}) \geq 0$.

3 A VERIFICATION THEOREM

The Hamilton-Jacobi inequality, which we studied in Section 1.7, is given, for a C^1 function $W(t,\hat{y})$, by:

$$\begin{aligned} &W_t(t,\hat{y}) + \langle W_{\hat{y}}(t,\hat{y}) \, , \, \hat{v}\rangle - \hat{L}(t,\hat{y},\hat{v}) \\ &\qquad \leq W_t(t,\hat{x}(t)) + \langle W_{\hat{y}}(t,\hat{x}(t)) \, , \, \dot{\hat{x}}(t)\rangle - \hat{L}(t,\hat{x}(t),\dot{\hat{x}}(t)) \end{aligned} \tag{HJI}$$

For our problem, since the constraints are not fixed endpoints, this inequality is not enough for sufficiency. In the following proposition, which is readily verified, we show what else is needed.

4. **Proposition:** Suppose $\hat{x}$ belongs to $\hat{X}_\bullet$ and there exist $\varepsilon > 0$ and a function W of class C^1 mapping $\hat{T}_0(\hat{x};\varepsilon)$ to R such that:

a. (HJI) holds for all $(t,\hat{y},\hat{v}) \in \hat{T}_1(\hat{x};\varepsilon)$.

b. $f(\hat{x}(\theta)) \leqslant f(\hat{y}(\theta))$ for all $\hat{y} \in \hat{X}_\bullet$ with $(t,\hat{y}(t),\dot{\hat{y}}(t)) \in \hat{T}_1(\hat{x};\varepsilon)$ where $f(\hat{y}) := W(\theta,\hat{y}) - W(0,c(\hat{y}))$.

Then $\hat{x}$ is a weak minimum for $(\hat{P})$. If (HJI) holds for all $(t,\hat{y},\hat{v})$ in $\hat{T}_0(\hat{x};\varepsilon) \times R^{np}$ and $f(\hat{x}(\theta)) \leqslant f(\hat{y}(\theta))$ for all $\hat{y} \in \hat{X}_\bullet$ with $(t,\hat{y}(t))$ in $\hat{T}_0(\hat{x};\varepsilon)$ then $\hat{x}$ is a strong minimum for $(\hat{P})$.

In Section 1.7 it is shown how the classical sufficient conditions for the basic problem in the calculus of variations imply precisely the existence of a C^1 function W satisfying (HJI). The proof we gave, based on that of F. H. Clarke and V. Zeidan, exhibits a verification function in terms of a solution of the Jacobi equation. It is given explicitly by:

$$W(t,\hat{y}) = \langle \hat{L}_{\hat{v}}(t,\hat{x}(t),\dot{\hat{x}}(t)) \ , \ \hat{y}\rangle + \frac{1}{2} \langle \hat{y} - \hat{x}(t) \ , \ V(t)(\hat{y} - \hat{x}(t))\rangle$$

where $V(t) = Q(t)Y^{-1}(t)$ and (Y,Q) is a solution, with $|Y(t)| \neq 0$ and $Y^*(t)Q(t) = Q^*(t)Y(t)$ for all $t \in [0,\theta]$, of the linear differential system:

$$\dot{Y}(t) = \hat{A}(t)Y(t) + \hat{B}(t)Q(t)$$

$$\dot{Q}(t) = (\hat{C} - \delta\hat{I})(t)Y(t) - \hat{A}^*(t)Q(t).$$

The constant δ is strictly positive, and the matrices $\hat{A}$, $\hat{B}$ and $\hat{C}$ are given by:

$$\hat{A}(t) = - \hat{L}_{\hat{v}\hat{v}}{}^{-1}(t)\ \hat{L}_{\hat{v}\hat{x}}(t)$$

$$\hat{B}(t) = \hat{L}_{\hat{v}\hat{v}}{}^{-1}(t)$$

$$\hat{C}(t) = \hat{L}_{\hat{x}\hat{x}}(t) - \hat{L}_{\hat{x}\hat{v}}(t)\ \hat{L}_{\hat{v}\hat{v}}{}^{-1}(t)\ \hat{L}_{\hat{v}\hat{x}}(t)$$

where, for example, $\hat{L}_{\dot{x}\dot{x}}(t)$ is an abbreviation of $\hat{L}_{\dot{x}\dot{x}}(t,\hat{x}(t),\dot{\hat{x}}(t))$. The matrix $V(t) = Q(t)Y^{-1}(t)$ is symmetric and satisfies on $[0,\theta]$ the matrix Riccati inequality:

$$\dot{V}(t) + V(t)\hat{A}(t) + \hat{A}^{*}(t)V(t) + V(t)\hat{B}(t)V(t) - \hat{C}(t) < 0. \tag{3.1}$$

Now, observe that, given $\hat{y} \in \hat{X}_{\bullet}$,

$$\langle \hat{L}_{\dot{v}}(\theta) , \hat{y}(\theta)\rangle - \langle \hat{L}_{\dot{v}}(0) , \hat{y}(0)\rangle = \langle L_{\dot{x}}(\tilde{x}(t_0+(N+1)\theta)) , \hat{y}_N(\theta)\rangle - \langle \hat{L}_{\dot{v}_0}(0) , \hat{y}_0(0)\rangle.$$

Hence, if we define, for all $\hat{y}$ in R^{np},

$$F(\hat{y}) := \langle \hat{y} - \hat{x}(\theta) , V(\theta)(\hat{y} - \hat{x}(\theta))\rangle - \langle c(\hat{y}) - \hat{x}(0) , V(0)(c(\hat{y}) - \hat{x}(0))\rangle$$

it follows that, for all $\hat{y} \in \hat{X}_{\bullet}$, $f(\hat{x}(\theta)) \leqslant f(\hat{y}(\theta))$ if, and only if, $F(\hat{y}(\theta)) \geqslant 0$. Consequently, we obtain sufficient conditions for problem $(\hat{P})$ (and hence, for the original one delay problem (P)), imposing the usual sufficient conditions for problem $(\hat{P}')$ together with the nonnegativity of F along the end points $\hat{y}(\theta)$, where $\hat{y}$ belongs to $\hat{X}_{\bullet}$ and, for a strong (weak) minimum, $\hat{y}$ lies on a tube (restricted tube) about $\hat{x}$. Of course, both for a weak and strong minima, this last condition for F holds if, for some $\varepsilon > 0$,

$$F(\hat{y}) \geqslant 0 \text{ for all } \hat{y} \in R^{nN} \times \{\xi\} \text{ with } |\hat{y} - \hat{x}(\theta)| < \varepsilon.$$

Now, let $Z := \{\hat{z} = (\hat{z}_0,\ldots,\hat{z}_N) \in R^{np} \mid \hat{z}_N = 0\}$ and observe that this condition for F holds if, for all $\hat{z} \in Z$,

i. $\langle F'(\hat{x}(\theta)) , \hat{z}\rangle = 0$

ii. $\langle \hat{z} , F''(\hat{x}(\theta))\hat{z}\rangle > 0$.

Condition (i) holds in view of the constraint $c(\hat{x}(\theta)) = \hat{x}(0)$ and (ii) is equivalent to $\langle \hat{z} , (V(\theta) - V(0)J)\hat{z}\rangle > 0$ where the constant matrix $J = c'(\hat{x}(\theta))c'(\hat{x}(\theta))$ is given given by $J_{ij} = I$ if $i + j = 2$ and $J_{ij} = 0$ otherwise. We shall label this condition (D1):

D1: $\langle \hat{z} , (V(\theta) - V(0)J)\hat{z}\rangle > 0$ for all $\hat{z} \in Z$, where $V(t)$ is a symmetric solution on $[0,\theta]$ of the matrix Riccati inequality (3.1).

Let us turn now to sufficiency through fields of extremals and see its relation with condition (D1). The verification theorem for $(\hat{P})$ in terms of the Hamilton-Jacobi partial differential equation is stated in the following proposition. Recall that the Hamiltonian for $\hat{L}$ is given, for all $(t,\hat{x},\hat{p})$ in $[0,\theta] \times R^{2np}$, by:

$$H(t,\hat{x},\hat{p}) := \sup\{\langle\hat{p},\hat{u}\rangle - \hat{L}(t,\hat{x},\hat{u}) \mid \hat{u} \in R^{np}\}.$$

5. **Proposition:** Suppose $\hat{x}$ belongs to $\hat{X}_e$ and there exist $\varepsilon > 0$ and a function W of class C^1 mapping $\hat{T}_0(\hat{x};\varepsilon)$ to R such that:

a. For all $(t,\hat{y}) \in \hat{T}_0(\hat{x};\varepsilon)$, $W_t(t,\hat{y}) + H(t,\hat{y},W_{\hat{y}}(t,\hat{y})) = 0$.

b. For all t in $[0,\theta]$,

$$H(t,\hat{x}(t),W_{\hat{y}}(t,\hat{x}(t))) = \langle W_{\hat{y}}(t,\hat{x}(t)) , \dot{\hat{x}}(t)\rangle - \hat{L}(t,\hat{x}(t),\dot{\hat{x}}(t)).$$

c. $f(\hat{x}(\theta)) \leqslant f(\hat{y}(\theta))$ for all $\hat{y} \in \hat{X}_e$ with $(t,\hat{y}(t)) \in \hat{T}_0(\hat{x};\varepsilon)$ where $f(\hat{y}) := W(\theta,\hat{y}) - W(0,c(\hat{y}))$.

Then $\hat{x}$ is a strong minimum for $(\hat{P})$.

As for (HJI), we know from Chapter 1 that the classical sufficient conditions for a strong minimum for $(\hat{P}')$ imply the existence of a C^1 function satisfying (5a) and (5b). The partial derivatives of W are given, for all $(t,\hat{y})$ in $\hat{T}_0(\hat{x};\varepsilon)$ by

$$W_t(t,\hat{y}) = \hat{L}(t,\hat{y},\Gamma(t,\hat{y})) - \langle P(t,\hat{y}) , \Gamma(t,\hat{y})\rangle$$

$$W_{\hat{y}}(t,\hat{y}) = \hat{L}_{\hat{v}}(t,\hat{y},\Gamma(t,\hat{y})) = P(t,\hat{y})$$

where $(\Gamma,\hat{T}_0(\hat{x};\varepsilon))$ is the Mayer-field constructed in Lemma 1.15 of which

$\hat{x}$ is an extremal. Recall that $\Gamma(t,\hat{y}) = \dot{\hat{x}}(t,\lambda(t,\hat{y}))$, where λ is the unique local solution of $\hat{y} = \hat{x}(t,\lambda(t,\hat{y}))$ and the family of extremals $\{\hat{x}(\cdot,\lambda)\}$ containing $\hat{x}$ are such that $|\hat{x}_\lambda(t,0)| \neq 0$ for all $t \in [0,\theta]$.

Now, condition (5c) is satisfied, as before, if, for all $\hat{z} \in Z$,

i. $\langle f'(\hat{x}(\theta)) , \hat{z}\rangle = 0$

ii. $\langle \hat{z} , f''(\hat{x}(\theta))\hat{z}\rangle > 0$.

Observe that $f'(\hat{x}(\theta)) = W_{\dot{\hat{y}}}(\theta,\hat{x}(\theta)) - W_{\dot{\hat{y}}}(0,\hat{x}(0))c'(\hat{x}(\theta))$ and so (i) holds in view of $c(\hat{x}(\theta)) = \hat{x}(0)$. For condition (ii), set $V(t) = Q(t)Y^{-1}(t)$ where

$$Q(t) := \hat{L}_{\hat{v}\dot{\hat{x}}}(t,\hat{x}(t),\dot{\hat{x}}(t))Y(t) + \hat{L}_{\hat{v}\hat{v}}(t,\hat{x}(t),\dot{\hat{x}}(t))\dot{Y}(t)$$

and $Y(t) := \hat{x}_\lambda(t,0)$. It follows that $f''(\hat{x}(\theta)) = V(\theta) - V(0)J$ and so (5c) holds if

D2: $\langle \hat{z} , (V(\theta) - V(0)J)\hat{z}\rangle > 0$ for all $\hat{z} \in Z$.

These facts, together with Theorem 1.4, imply the following sufficient conditions for (P):

6. **Theorem:** Suppose $\hat{L} \in C^2(\hat{x})$, $\hat{L}_{\hat{v}} \in C^1$ and $x \in X_\bullet$. Let $\hat{x} := \Phi(x)$ and suppose the following holds:

i. For all $\hat{y} \in \hat{Y}$, $\hat{I}'(\hat{x};\hat{y}) = 0$.

ii. For all $\hat{y} \in \hat{Y}$, $\hat{I}''(\hat{x};\hat{y}) > 0$.

iii. For all $t \in [0,\theta]$, $\hat{L}_{\hat{v}\hat{v}}(t,\hat{x}(t),\dot{\hat{x}}(t)) > 0$.

If any of conditions (D1) or (D2) is satisfied, then x is a weak minimum for (P). Moreover, if also $\hat{L}$ is regular or, for some $\varepsilon > 0$,

iv. $\hat{E}(t,\hat{y},\dot{\hat{y}},\hat{u}) \geqslant 0$ for all $(t,\hat{y},\dot{\hat{y}}) \in \hat{T}_1(\hat{x};\varepsilon)$ and $\hat{u} \in R^{np}$,

then x is a strong minimum for (P).

4 CHARACTERIZATIONS

We turn now to characterize the conditions derived in the previous two sections. Let us start evaluating the first and second partial derivatives of $\hat{L}$ and seeing their relation with those of the original Lagrangian. Let $t \in [0,\theta]$ and $\hat{x} = (\hat{x}_0,\ldots,\hat{x}_N)$ and $\hat{v} = (\hat{v}_0,\ldots,\hat{v}_N) \in R^{np}$. Directly from the definitions we have:

$$\hat{L}_{\hat{v}_k}(t,\hat{x},\hat{v}) = L_{\dot{x}}(t+t_0,\hat{x}_0,\phi(t+t_0-\theta),\hat{v}_0) \qquad k = 0$$

$$= L_{\dot{x}}(t+t_0+k\theta,\hat{x}_k,\hat{x}_{k-1},\hat{v}_k) \qquad k = 1,\ldots,N.$$

$$\hat{L}_{\hat{x}_k}(t,\hat{x},\hat{v}) = L_x(t+t_0,\hat{x}_0,\phi(t+t_0-\theta),\hat{v}_0) + L_u(t+t_0+\theta,\hat{x}_1,\hat{x}_0,\hat{v}_1) \qquad k = 0$$

$$= L_x(t+t_0+k\theta,\hat{x}_k,\hat{x}_{k-1},\hat{v}_k)$$
$$+ L_u(t+t_0+(k+1)\theta,\hat{x}_{k+1},\hat{x}_k,\hat{v}_{k+1}) \qquad k = 1,\ldots,N-1$$

$$= L_x(t+t_0+N\theta,\hat{x}_N,\hat{x}_{N-1},\hat{v}_N) \qquad k = N.$$

We shall find it useful to evaluate explicitly the first derivatives of $\hat{L}$ along $\Phi(x)$, where x is any trajectory in $X_\bullet$.

7. **Lemma:** Let $x \in X_\bullet$ and set $\hat{x} := \Phi(x)$. Then, for all t in $[0,\theta]$,

$$\hat{L}_{\hat{v}_k}(t,\hat{x}(t),\dot{\hat{x}}(t)) = L_{\dot{x}}(\tilde{x}(t+t_0+k\theta)) \qquad k = 0,\ldots,N$$

$$\hat{L}_{\hat{x}_k}(t,\hat{x}(t),\dot{\hat{x}}(t)) = L_x(\tilde{x}(t+t_0+k\theta)) + L_u(\tilde{x}(t+t_0+(k+1)\theta)) \qquad k = 0,\ldots,N-1$$

$$= L_x(\tilde{x}(t+t_0+N\theta)) \qquad k = N.$$

The following proposition shows that Euler's equation for the Lagrangians L and $\hat{L}$ are equivalent.

8. Proposition: Let $x \in X_\bullet$ and assume $L_{\dot{x}} \circ \tilde{x}$ is continuous and $L_u(\tilde{x}(t)) = 0$ for all t in $(t_1,t_2]$. Then the following are equivalent:

i. $I'(x;y) = 0$ for all y in Y.

ii. $\hat{I}'(\Phi(x);\hat{y}) = 0$ for all $\hat{y}$ in $\hat{Y}$.

Proof: Set $\hat{x} := \Phi(x)$. (i) $\Rightarrow$ (ii): By Proposition 2.6, x satisfies Euler's equation, i.e., there exists a constant c in R^n such that, for all t in $[t_0,t_1]$,

$$L_{\dot{x}}(\tilde{x}(t)) = P(t;x) + c$$

where $P(\cdot;x)$ is given by:

$$\begin{aligned} P(t;x) &= \int_{t_0}^{t} L_x(\tilde{x}(s))ds + \int_{t_0}^{t+\theta} L_u(\tilde{x}(s))ds \quad t_0 \leq t \leq t_1-\theta \\ &= \int_{t_0}^{t} L_x(\tilde{x}(s))ds + \int_{t_0}^{t_1} L_u(\tilde{x}(s))ds \quad t_1-\theta \leq t \leq t_1. \end{aligned}$$

Define $\hat{c} = (\hat{c}_0,\ldots,\hat{c}_N)$ as follows: for $k = 0,1,\ldots,N-1$

$$\hat{c}_k = \int_{t_0}^{t_0+k\theta} L_x(\tilde{x}(s))ds + \int_{t_0}^{t_0+(k+1)\theta} L_u(\tilde{x}(s))ds + c$$

$$\hat{c}_N = \int_{t_0}^{t_0+N\theta} L_x(\tilde{x}(s))ds + \int_{t_0}^{t_1} L_u(\tilde{x}(s))ds + c$$

In view of Lemma 7, for all t in $[0,\theta]$,

$$\begin{aligned} \hat{L}_{\hat{v}_k}(t,\hat{x}(t),\dot{\hat{x}}(t)) &= L_{\dot{x}}(\tilde{x}(t+t_0+k\theta)) \\ &= P(t+t_0+k\theta;x) + c \\ &= \int_0^t \hat{L}_{\hat{x}_k}(s,\hat{x}(s),\dot{\hat{x}}(s))ds + \hat{c}_k. \end{aligned}$$

Hence, $\hat{L}_{\hat{v}}(t,\hat{x}(t),\dot{\hat{x}}(t)) = \int_0^t \hat{L}_{\hat{x}}(s,\hat{x}(s),\dot{\hat{x}}(s))ds + \hat{c}$ and Proposition 1.5

implies (ii).

(ii) ⇒ (i): By Proposition 1.5, there exists a constant $\hat{c} = (\hat{c}_0,\ldots,\hat{c}_N)$ in R^{np} such that, for all t in $[0,\theta]$,

$$\hat{L}_{\dot{v}}(t,\hat{x}(t),\dot{\hat{x}}(t)) = \int_0^t \hat{L}_{\dot{x}}(s,\hat{x}(s),\dot{\hat{x}}(s))ds + \hat{c}$$

Applying Lemma 7 we obtain that, for $k = 0,\ldots,N-1$ and $t_0+k\theta \leqslant t \leqslant t_0+(k+1)\theta$,

$$L_{\dot{x}}(\tilde{x}(t)) = \int_{t_0+k\theta}^{t} \{L_x(\tilde{x}(s)) + L_u(\tilde{x}(s+\theta))\}ds + \hat{c}_k$$

and, for $t_0+N\theta \leqslant t \leqslant t_2$,

$$L_{\dot{x}}(\tilde{x}(t)) = \int_{t_0+N\theta}^{t} L_x(\tilde{x}(s))ds + \hat{c}_N$$

Since $L_{\dot{x}} \circ \tilde{x}$ is continuous we have, for $k = 1,\ldots,N-1$:

$$\hat{c}_k = L_{\dot{x}}(\tilde{x}(t_0+k\theta)) = \int_{t_0+(k-1)\theta}^{t_0+k\theta} \{L_x(\tilde{x}(s)) + L_u(\tilde{x}(s+\theta))\}ds + \hat{c}_{k-1}$$

For $k = N$ the integrand is replaced by $L_x(\tilde{x}(s))$. Consequently, for all t in $[t_0,t_1-\theta]$,

$$L_{\dot{x}}(\tilde{x}(t)) = \int_{t_0}^{t} \{L_x(\tilde{x}(s)) + L_u(\tilde{x}(s+\theta))\}ds + \hat{c}_0$$

and, for t in $[t_1-\theta,t_1]$,

$$L_{\dot{x}}(\tilde{x}(t))= \int_{t_0}^{t} L_x(\tilde{x}(s))ds + \hat{c}_0.$$

The required result follows now by Proposition 2.6.

Remark: It should be noted that, in view of Lemma 7, the condition of Weierstrass for L implies the corresponding condition for $\hat{L}$, but the converse is not necessarily true.

We turn now to characterize the positivity of the second variation. For the second partial derivatives of $\hat{L}$ we have:

$$\hat{L}_{\hat{v}_j\hat{v}_k}(t,\hat{x},\hat{v}) = L_{\dot{x}\dot{x}}(t+t_0,\hat{x}_0,\phi(t+t_0-\theta),\hat{v}_0) \qquad j = k = 0$$

$$= L_{\dot{x}\dot{x}}(t+t_0+k\theta,\hat{x}_k,\hat{x}_{k-1},\hat{v}_k) \qquad j = k \neq 0$$

$$= 0 \quad \text{otherwise}$$

$$\hat{L}_{\hat{v}_j\hat{x}_k}(t,\hat{x},\hat{v}) = L_{\dot{x}x}(t+t_0,\hat{x}_0,\phi(t+t_0-\theta),\hat{v}_0) \qquad j = k = 0$$

$$= L_{\dot{x}x}(t+t_0+k\theta,\hat{x}_k,\hat{x}_{k-1},\hat{v}_k) \qquad j = k \neq 0$$

$$= L_{\dot{x}u}(t+t_0+j\theta,\hat{x}_j,\hat{x}_{j-1},\hat{v}_j) \qquad k = j - 1$$

$$= 0 \quad \text{otherwise}$$

$$\hat{L}_{\hat{x}_j\hat{v}_k}(t,\hat{x},\hat{v}) = L_{x\dot{x}}(t+t_0,\hat{x}_0,\phi(t+t_0-\theta),\hat{v}_0) \qquad j = k = 0$$

$$= L_{x\dot{x}}(t+t_0+k\theta,\hat{x}_k,\hat{x}_{k-1},\hat{v}_k) \qquad j = k \neq 0$$

$$= L_{u\dot{x}}(t+t_0+j\theta,\hat{x}_j,\hat{x}_{j-1},\hat{v}_j) \qquad k = j + 1$$

$$= 0 \quad \text{otherwise}$$

$$\hat{L}_{\hat{x}_j\hat{x}_k}(t,\hat{x},\hat{v}) = L_{xx}(t+t_0,\hat{x}_0,\phi(t+t_0-\theta),\hat{v}_0) + L_{uu}(t+t_0+\theta,\hat{x}_1,\hat{x}_0,\hat{v}_1) \qquad j = k = 0$$

$$= L_{xx}(t+t_0+k\theta,\hat{x}_k,\hat{x}_{k-1},\hat{v}_k) + L_{uu}(t+t_0+(k+1)\theta,\hat{x}_{k+1},\hat{x}_k,\hat{v}_{k+1})$$

$$j = k \in \{1,\ldots,N-1\}$$

$$= L_{xx}(t+t_0+N\theta,\hat{x}_N,\hat{x}_{N-1},\hat{v}_N) \qquad k = j = N$$

$$= L_{ux}(t+t_0+k\theta,\hat{x}_k,\hat{x}_{k-1},\hat{v}_k) \qquad k = j + 1$$

$$= L_{xu}(t+t_0+j\theta,\hat{x}_j,\hat{x}_{j-1},\hat{v}_j) \qquad k = j - 1$$

$$= 0 \quad \text{otherwise}$$

9. **Lemma:** Let $x \in X_\bullet$, $\hat{y} \in \hat{X}$ and $\hat{x} := \Phi(x)$. Then, for all $t \in [0,\theta]$,

$$[\hat{L}_{\hat{v}\hat{v}}(t)\dot{\hat{y}}(t)]_k = L_{\dot{x}\dot{x}}(\tilde{x}(t+t_0+k\theta))\dot{\hat{y}}_k(t) \qquad k = 0,1,\ldots,N$$

$$[\hat{L}_{\hat{v}\hat{x}}(t)\hat{y}(t)]_k = L_{\dot{x}x}(\tilde{x}(t+t_0))\hat{y}_0(t) \qquad k = 0$$

$$= L_{\dot{x}u}(\tilde{x}(t+t_0+k\theta))\hat{y}_{k-1}(t) + L_{\dot{x}x}(\tilde{x}(t+t_0+k\theta))\hat{y}_k(t) \quad k = 1,\ldots,N$$

$$[\hat{L}_{\hat{x}\hat{v}}(t)\dot{\hat{y}}(t)]_k = L_{x\dot{x}}(\tilde{x}(t+t_0+k\theta))\dot{\hat{y}}_k(t) + L_{u\dot{x}}(\tilde{x}(t+t_0+(k+1)\theta))\dot{\hat{y}}_{k+1}(t) \qquad k = 0,1,\ldots,N-1$$

$$= L_{x\dot{x}}(\tilde{x}(t+t_0+N\theta))\dot{\hat{y}}_N(t) \qquad k = N$$

$$[\hat{L}_{\hat{x}\hat{x}}(t)\hat{y}(t)]_k = [L_{xx}(\tilde{x}(t+t_0)) + L_{uu}(\tilde{x}(t+t_0+\theta))]\hat{y}_0(t) + L_{ux}(\tilde{x}(t+t_0+\theta))\hat{y}_1(t) \qquad k = 0$$

$$= L_{xu}(\tilde{x}(t+t_0+k\theta))\hat{y}_{k-1}(t) + L_{ux}(\tilde{x}(t+t_0+(k+1)\theta))\hat{y}_{k+1}(t) + [L_{xx}(\tilde{x}(t+t_0+k\theta)) + L_{uu}(\tilde{x}(t+t_0+(k+1)\theta))]\hat{y}_k(t) \qquad k = 1,\ldots,N-1$$

$$= L_{xu}(\tilde{x}(t+t_0+N\theta))\hat{y}_{N-1}(t) + L_{xx}(\tilde{x}(t+t_0+N\theta))\hat{y}_N(t) \qquad k = N.$$

Comparing Propositions 1.9 and 2.11, it is clear by Lemma 9 that, if $t_1 = t_2$, then the Jacobi equation for (P) coincides with the Jacobi equation for $(\hat{P})$. Based on this fact, we define the notion of conjugate sequence as follows:

10. Definition: Given $x \in X_*$, we shall say that a sequence of points $\{s+t_0+k\theta \mid k = 0,\ldots,N\}$, with $s \in (0,\theta]$, is **conjugate to t_0 with respect to x**, if there exists $y \in X$ satisfying (JE) and such that:

i. $y(t) = 0$ for all $t \in [t_0-\theta,t_0]$

ii. $y(t+t_0+k\theta)$ is not identically zero for $t \in (0,s)$ $k = 0,\ldots,N$.

iii. $y(t_0+k\theta) = y(s+t_0+k\theta) = 0$ $k = 0,\ldots,N$.

We redefine the condition of Jacobi as follows: x will be said to satisfy **Jacobi's condition** if there are no conjugate sequences to t_0 with respect to x on (t_0,t_1) and **Jacobi's strengthened condition** if there are no conjugate sequences to t_0 with respect to x on $(t_0,t_1]$.

Now, in view of Lemma 9, given $x \in X_*$ and $y \in X$, if $t_1 = t_2$ then

$$I''(x;y) = \hat{I}''(\Phi(x);\Phi(y))$$

and so $I''(x;y) > 0$ for all y in $Y - \{0\}$ implies $\hat{I}''(\Phi(x);\hat{y}) > 0$ for all $\hat{y}$ in $\hat{Y} - \{0\}$. The converse is not necessarily true, but Proposition 1.12 suggests a condition, similar to (D1) and (D2), which implies the strict positivity of $I''(x;\cdot)$ along $Y - \{0\}$. Assume $L \in C^2(x;u)$, $L_{\dot{x}} \in C^1$, $t_1 = t_2$ and x is a nonsingular extremal satisfying the Legendre condition. It follows from the proof of Proposition 1.12 that $I''(x;y) > 0$ for all y in $Y - \{0\}$ if, and only if, there exists (Y,Q), a matrix solution of the Jacobi equation for $\hat{L}$ (with respect to $\Phi(x)$), with $|Y(t)| \neq 0$ and $Y^*(t)Q(t) = Q^*(t)Y(t)$ for all $t \in [0,\theta]$ such that, for all $y \in Y - \{0\}$,

$$\int_0^\theta \langle\hat{z}(t)\ ,\ \hat{L}_{\dot{\hat{v}}\dot{\hat{v}}}(t,\hat{x}(t),\dot{\hat{x}}(t))\hat{z}(t)\rangle dt + R(y) > 0$$

where

$$R(y) = \langle\hat{y}(\theta)\ ,\ V(\theta)\hat{y}(\theta)\rangle - \langle\hat{y}(0)\ ,\ V(0)\hat{y}(0)\rangle$$

$V(t) = Q(t)Y^{-1}(t)$, $\hat{y} = \Phi(y)$ and $\hat{z}(t) = \dot{\hat{y}}(t) + Y(t)\dot{Y}^{-1}(t)\hat{y}(t)$.

Consequently, if there exists (Y,Q) satisfying the conditions above and $\langle \hat{z} , (V(\theta) - V(0)J)\hat{z}\rangle > 0$ for all $\hat{z} \in Z$, then R(y) will be strictly positive unless $y \equiv 0$, implying $I''(x,y) > 0$ for all y in Y - {0}. The matrix $V(t) = Q(t)Y^{-1}(t)$ is symmetric and, since $V(t)\dot{Y}(t) + \dot{V}(t)Y(t) = \dot{Q}(t)$, it follows that V(t) satisfies the matrix Riccati equation:

$$\dot{V}(t) + V(t)\hat{A}(t) + \hat{A}^*(t)V(t) + V(t)\hat{B}(t)V(t) - \hat{C}(t) = 0. \qquad (3.2)$$

The converse is also true, that is, if there exists a symmetric matrix V(t) satisfying (3.2) and $\langle \hat{z} , (V(\theta) - V(0)J)\hat{z}\rangle > 0$ for all $\hat{z} \in Z$, then $I''(x;y) > 0$ for all y in Y - {0}. This follows by direct calculation since, for all $y \in X$,

$$\hat{I}''(\Phi(x);\hat{y}) = \int_0^\theta \langle \hat{z}(t) , \hat{L}_{\dot{\hat{v}}\dot{\hat{v}}}(t,\hat{x}(t),\dot{\hat{x}}(t))\hat{z}(t)\rangle dt + R(y)$$

where R and $\hat{y}$ are defined as above, and

$$\hat{z}(t) = \dot{\hat{y}}(t) + \hat{L}_{\dot{\hat{v}}\dot{\hat{v}}}(t)(\hat{L}_{\dot{\hat{v}}\hat{x}}(t) - V(t))\hat{y}(t).$$

We shall label this last condition (D3):

D3: There exists V(t), a symmetric matrix solution of (3.2), such that $\langle \hat{z} , (V(\theta) - V(0)J)\hat{z}\rangle > 0$ for all $\hat{z} \in Z$.

Observe that (D2) implies (D3) and so we can reduce the three conditions into one as follows: we shall say that x satisfies **Riccati's condition** if there exists V(t), a symmetric matrix solution of (3.1) or (3.2), such that $\langle \hat{z} , (V(\theta) - V(0)J)\hat{z}\rangle > 0$ for all $\hat{z} \in Z$.

We are now in a position to state the necessary and sufficient conditions for (P) obtained so far.

11. Theorem: Suppose $L \in C^2(x,u,\dot{x})$, $t_1 = t_2$ and x is a trajectory solving (P). Then x satisfies Euler's equation and the conditions of Legendre and Weierstrass. If also x is nonsingular then x satisfies Jacobi's condition.

12. **Theorem:** Suppose $L \in C^2(x,u)$, $L_{\dot{x}} \in C^1$, $t_1 = t_2$, and $x \in X_\bullet$ is a nonsingular trajectory satisfying Euler's equation, Legendre's and Riccati's conditions and the strengthened condition of Jacobi. Then x is a weak minimum for (P). If also L is regular on A or $\Phi(x)$ satisfies the strengthened condition of Weierstrass (for $\hat{L}$), then x is a strong minimum for (P).

5 EXAMPLES

5.1 A problem with no minimum.

Let $L(t,x,u,\dot{x}) = \dot{x}^2 - x^2 - u$ and suppose we are given a point $\xi \in R$, $\phi:[-\pi,0] \to R$ and (P), the problem under consideration, is to minimize

$$I(x) = \int_0^{2\pi} [\dot{x}(t)^2 - x(t)^2 - x(t-\pi)]dt$$

over all piecewise-C^1 trajectories $x:[-\pi,2\pi] \to R$ satisfying $x(t) = \phi(t)$ on $[-\pi,0]$ and $x(2\pi) = \xi$.

Suppose x is any piecewise-C^1 trajectory. The way L is defined implies that x is nonsingular. The Jacobi equation, relative to x, is given by:

$$\frac{d}{dt}[2\dot{y}(t)] = -2y(t) \qquad 0 \leqslant t \leqslant 2\pi$$

Let $y(t) = \sin(t)$ for all $t \in [0,2\pi]$ and $y(t) = 0$ for $t \in [-\pi,0]$. This trajectory solves the Jacobi equation, it is nonvanishing on $(0,\pi)$, $y(t) = 0$ for all $t \in [-\pi,0]$, $y(\pi) = 0$ and

$$\int_{\pi}^{2\pi} \langle y(t-\theta) , C^*(t)\dot{y}(t) + E^*(t)y(t)\rangle dt = 0.$$

This implies that $\pi \in (0,2\pi)$ is conjugate to 0 with respect to x. By Theorem 2.21, x cannot be a minimum for (P).

5.2 A weak minimum which is not a strong minimum.

Let $L(t,x,u,\dot{x}) = \dot{x}^2 - 4x\dot{x}^3 - 2u\dot{x}^3 + 2t\dot{x}^4$,

$\theta = 1/2$, $\phi(t) = 0$ for $-1/2 \leqslant t \leqslant 0$, $t_0 = 0$, $t_1 = 1$ and $\xi = 0$

and consider the problem:

$$P: \quad \text{Minimize} \quad \int_0^1 \{\dot{x}(t)^2 - 4x(t)\dot{x}(t)^3 - 2x(t-\theta)\dot{x}(t)^3 + 2t\dot{x}(t)^4\}dt$$

subject to $x(t) = \phi(t)$ on $[-1/2,0]$ and $x(1) = \xi$.

Consider the trajectory $x_0(t) \equiv 0$. We shall show that x_0 is a weak minimum for (P).

i. x_0 satisfies Euler's equation: from the definition of L, Euler's equation is given by:

$$\frac{d}{dt} \{2\dot{x}(t) - 12x(t)\dot{x}(t)^2 - 6x(t-\theta)\dot{x}(t)^2 + 8t\dot{x}(t)^3\}$$

$$= -4\dot{x}(t)^3 - 2\dot{x}(t+\theta)^3 \quad 0 \leqslant t \leqslant 1/2$$

$$= -4\dot{x}(t)^3 \quad 1/2 \leqslant t \leqslant 1$$

which is satisfied by x_0.

ii. x_0 is nonsingular: $L_{\dot{x}\dot{x}}(\tilde{x}_0(t)) = 2$ for all $t \in [0,1]$.

iii. $I''(x_0;y) > 0$ for all $y \in Y - \{0\}$: this follows since

$$I''(x_0;y) = \int_0^1 2\dot{y}(t)^2 dt$$

By Theorem 2.5, it follows that x_0 is a weak minimum for (P).

Now, the trajectory x_0 satisfies the condition of Weierstrass, since

$$E(t,x_0(t),x_0(t-\theta),\dot{x}_0(t),v) = v^2 + 2tv^4 \geqslant 0$$

but, as we shall see below, x_0 is not a strong minimum for (P). Define for k and h strictly positive:

$$\begin{aligned} x(t) &= 0 \qquad -1/2 \leqslant t \leqslant 0 \\ &= kt/h \qquad 0 \leqslant t \leqslant h \\ &= k(1-t)/(1-h) \qquad h \leqslant t \leqslant 1 \end{aligned}$$

Evaluating I along x, and assuming $h \leqslant 1/2$, we obtain:

$$\begin{aligned} I(x) = &\int_0^h \{k^2/h^2 - 4tk^4/h^4 + 2tk^4/h^4\}dt \\ &+ \int_h^1 \{k^2/(1-h)^2 - 4(1-t)k^4/(1-h)^4 + 2tk^4/(1-h)^4\}dt \\ &+ \int_\pi^{\pi+h} 2(t-\theta)k^4/h(1-h)^3\,dt + \int_{\pi+h}^1 2(1-t+\theta)k^4/(1-h)^4\,dt \\ = &-k^4/h^2 + k^2/h + k^2/(1-h) + k^4(3-h)/(1-h)^3 \\ &+ k^4h/(1-h)^3 + k^4(3/4 - 2h + h^2)/(1-h)^4 \end{aligned}$$

We have $|x(t)| \leqslant k$ for $-1/2 \leqslant t \leqslant 1$ and, for each fixed k, $I(x) < 0$ whenever h is sufficiently small. Since x satisfies the endpoint constraints, x_0 does not afford a strong minimum to I.

Corrigenda

The proofs of the results in Chapter 3 are all correct only under the assumption that the time interval is commensurate with the time delay, i.e.

$$t_1 = t_2.$$

PART II

GENERIC CONDITIONS FOR NORMALITY IN OPTIMAL CONTROL THEORY

CHAPTER 4

NORMALITY CONDITIONS

1 INTRODUCTION

There are optimal control problems where the necessary conditions of Pontryagin's maximum principle do not involve the cost or performance index, adding no information to the one already specified by the dynamics, control set and end conditions. These problems are called "abnormal" and several attempts to characterize them have been made. The question has remained unanswered since the maximum principle was established. The strongest result, due to R. B. Vinter [30], states that, under mild conditions (satisfied by the so-called "relaxed problem"), if one perturbs the original problem through translations of the endpoint set, there is a dense set of problems where normality can be guaranteed.

In this chapter we extend this result based on a "lower semilipschitz" property of the value function known as "calmness". When the equations of motion are linear in the state variable, we enlarge the previous dense set to a full (Lebesgue) measure set. In other words, we show that normality holds for all problems obtained by translating the original endpoint set, except possibly in directions that belong to a set of measure zero.

We also introduce a different way of perturbing the endpoint set (enlarging or diminishing instead of translating it) and prove that, dealing with this kind of perturbations, one obtains normality almost everywhere even if the equations of motion are not linear in the state variable.

2 NORMALITY IN A MATHEMATICAL PROGRAMMING PROBLEM

Dealing with a mathematical programming problem such as:

$$P: \begin{cases} \text{minimize } f(x) \text{ subject to:} \\ g_i(x) \leqslant 0 \quad (i \in I = \{1,\dots,m\}) \end{cases}$$

where f, $g_i : R^n \to R$ are of class C^1 $(i \in I)$, one is concerned with the linearly independence of

$$\{\nabla g_i(x) \mid i \in I(x)\} \tag{4.1}$$

where $I(x) = \{i \in I \mid g_i(x) = 0\}$, in order to have the Lagrange multiplier rule as a necessary condition for local optimality. That is, if x is a local solution of (P) and (4.1) is linearly independent, then there exist $\lambda_1,\dots,\lambda_m \geqslant 0$, with $\lambda_i = 0$ if $i \notin I(x)$, such that

$$F := f + \sum_1^m \lambda_i g_i \quad \Rightarrow \quad \nabla F(x) = 0.$$

In 1979, J. Spingarn and R. Rockafellar [28] proved that, if $g_1,\dots,g_m$ are of class C^n, then "most" problems satisfy these conditions, in the sense that, for all $u \in R^m$, except in a set of measure zero, if x is any feasible solution of

$$P_u: \begin{cases} \text{minimize } f(x) \text{ subject to:} \\ g_i(x) \leqslant u_i \quad (i \in I) \end{cases}$$

and $I(x,u) = \{i \in I \mid g_i(x) = u_i\}$, then $\{\nabla g_i(x) \mid i \in I(x,u)\}$ is linearly independent. This result helps to show that, with the same kind of perturbations, most problems are such that every local solution satisfies the strong second-order conditions for optimality.

A similar result was proved in 1976 by F. Clarke [7] assuming merely local Lipschitz continuity on the functions delimiting the problem. It is based on a theorem of I. Ekeland, the theory of generalized gradients and the concept of calmness. The central result is the following:

If $(X, \| \ \|)$ is any Banach space, $f, g_i : X \to R$ are locally Lipschitz $(i \in I)$ and x^* is a local solution of (P), then there exist λ_0, λ_i $(i \in I)$ not all zero, such that:

i. $\lambda_0, \lambda_i \geqslant 0 \quad (i \in I)$

ii. $\lambda_i g_i(x^*) = 0 \quad (i \in I)$

iii. $0 \in \lambda_0 \partial f(x^*) + \sum_1^m \lambda_i \partial g_i(x^*)$,

where "∂" denotes the generalized gradient.

The problem (P) is called "normal" if, whenever x^* solves (P), there exist λ_0, λ_i $(i \in I)$ satisfying (i)-(iii) with $\lambda_0 > 0$. Given $u \in R^m$, (P_u) is called "calm" if $\phi(u) \in R$ and

$$\liminf_{v \to u} \frac{\phi(v) - \phi(u)}{|v - u|} > -\infty$$

where $\phi : R^m \to [-\infty, \infty]$ is given by $\phi(u) = \inf\{f(x) \mid g_i(x) \leqslant u_i,\ i \in I\}$ and $|\cdot|$ denotes the norm in R^m.

It is not difficult to show that, if (P_u) is calm and x^* a solution of (P_u), there exist $K, \varepsilon > 0$ such that x^* solves:

$$L: \quad \begin{array}{l} \text{minimize } f(x) + K \sum_1^m \max\{0\ ,\ g_i(x) - u_i\} \\ \text{subject to: } \|x - x^*\| < \varepsilon \end{array}$$

which in turn implies that (P_u) is normal. This implies that, if U is a neighborhood of $0 \in R^m$ where ϕ is finite then, for all $u \in U$ except in a set of measure zero, (P_u) is normal, for ϕ is decreasing as a function of each component of u separately, so differentiable a.e., and so (P_u) calm for almost all $u \in U$.

3 STATEMENT OF THE PROBLEM

We shall be concerned with the (Mayer) optimal control problem:

$$P: \begin{cases} \text{minimize } g(x(1)) \text{ subject to:} \\ x \in X_{\bullet}, \quad x(1) \in C \end{cases}$$

where:

$$X_{\bullet} = \{x \in X \mid x(0) \in C_0 \text{ and there exists } u:[0,1] \to R^m \text{ such that } u(t) \in U(t) \text{ and } \dot{x}(t) = f(t,x(t),u(t)) \text{ almost everywhere}\}$$

$$X = \{x:[0,1] \to R^n \mid x \text{ is absolutely continuous}\}$$

and the following is satisfied:

i. $C, C_0 \subset R^n$ are closed.

ii. g is locally Lipschitz.

iii. For each $x \in R^n$, $f(\cdot,x,\cdot)$ is $L \times B$ measurable, where $L \times B$ denotes the σ-algebra of subsets of $[0,1] \times R^m$ generated by product sets $M \times N$, where $M \subset [0,1]$ is Lebesgue measurable and $N \subset R^m$ a Borel set.

iv. For each $x \in X$ there exist $k \in L^1(0,1)$ and Ω, an ε-neighborhood of x, such that, for all $t \in [0,1]$ and $u \in U(t)$, $f(t,\cdot,u)$ is Lipschitz on Ω of rank $k(t)$.

v. $Gr(U):=\{(t,u) \in [0,1] \times R^m \mid u \in U(t)\}$ is $L \times B$ measurable.

Under these hypotheses, the maximum principle states the following (we refer to [9]):

1. **Theorem:** If x^* solves (P) with u^* corresponding control, then there exist $\lambda \in \{0,1\}$ and an absolutely continuous function p mapping [0,1] to $\mathbb{R}^n$, not both zero, such that:

a. $-\dot{p}(t) \in p(t)\partial_x f(t,x^*(t),u^*(t))$ a.e.

b. $\langle p(t), f(t,x^*(t),u^*(t))\rangle = \sup\{\langle p(t), f(t,x^*(t),u)\rangle \mid u \in U(t)\}$ a.e.

c. $p(0) \in N_{C_0}(x^*(0))$

d. $-p(1) \in N_C(x^*(1)) + \lambda\partial g(x^*(1))$

(the notation $\partial_x f$ refers to the generalized Jacobian of $x \to f(t,x,u^*(t))$ and $N_C(x)$ denotes the normal cone to C at $x \in C$).

The problem we shall be dealing with is that of giving conditions that guarantee the nonvanishing of the cost's multiplier. In other words, under what conditions can λ can be taken to be 1.

4 CALMNESS AND NORMALITY

For the optimal control problem, we shall call (P) "normal" if, whenever x^* solves (P), there exist $(p(\cdot),\lambda)$ satisfying the maximum principle with $\lambda = 1$. Before defining the concept of "calmness", we start giving a simple condition that implies normality:

2. **Lemma:** Suppose that, whenever x^* solves (P), there exist K and ε positive, such that x^* solves:

$$L: \quad \begin{array}{l} \text{minimize} \quad g(x(1)) + K\,\text{dist}(x(1),C) \\ \text{subject to:} \quad x \in X_\bullet \text{ and } \|x - x^*\| < \varepsilon \end{array}$$

where $\|\cdot\|$ denotes the sup norm in X. Then (P) is normal.

Proof: Suppose x^* solves (P). Let K and $\varepsilon > 0$ be such that x^* solves L and apply the maximum principle for this problem. We obtain $(p(\cdot),\lambda)$ satisfying (a), (b) and (c) exactly as for problem (P), and (d) becomes:

$$-p(1) \in \lambda\,[\partial g(x^*(1)) + K\,\partial\text{dist}(x^*(1),C)].$$

Now, if λ were 0, we would have $-p(1) = 0$ which, together with (a), implies $p \equiv 0$. The contradiction implies $\lambda = 1$ and so (d) results as for problem (P).

As we shall see below, the notion of calmness implies the assumption of Lemma 2. Observe first what happens if we assume the contrary, that is, there exist sequences $\varepsilon_i \downarrow 0$, $K_i \uparrow \infty$ and $x_i \in X_\bullet$ satisfying $\|x_i - x^*\| < \varepsilon_i$, such that, for all $i \in \mathbf{N}$,

$$g(x_i(1)) - g(x^*(1)) < -K_i\text{dist}(x_i(1),C).$$

Since C is closed, for all $i \in \mathbf{N}$ there exists $c_i \in C$ such that

$$|x_i(1) - c_i| = \text{dist}(x_i(1),C).$$

Set $\alpha_i = x_i(1) - c_i$ and define the "value function" for all $a \in \mathbf{R}^n$:

$$\phi(a) := \inf\{g(x(1)) \mid x \in X_\bullet \text{ and } x(1) \in C + \{a\}\}.$$

We clearly obtain that

$$\frac{\phi(\alpha_i) - \phi(0)}{|\alpha_i|} \to -\infty, \; i \to \infty.$$

So, if we consider the family of problems, for $a \in \mathbf{R}^n$:

$$P_a: \begin{cases} \text{minimize } g(x(1)) \text{ subject to:} \\ x \in X_\bullet, \quad x(1) \in C + \{a\} \end{cases}$$

and call (P_a) "calm" if $\phi(a) \in R$ and

$$\liminf_{b \to a} \frac{\phi(b) - \phi(a)}{|b - a|} > -\infty,$$

then the above argument and Lemma 2 show that calmness implies normality.

This argument also induces in a natural way a different kind of perturbations to the original problem that gives generic conditions for normality.

3. **Theorem:** Consider the following family of problems for $a \in R^n$:

$$Q_a: \begin{cases} \text{minimize } g(x(1)) \text{ subject to:} \\ x \in X_\bullet, \quad \text{dist}(x(1),C) \leq |a| \end{cases}$$

Then, for almost all $a \in R^n$, Q_a is normal.

Proof: Define for all $a \in R^n$:

$$\psi(a) := \inf\{g(x(1)) \mid x \in X_\bullet \text{ and } \text{dist}(x(1),C) \leq |a|\}.$$

Now, the same argument as above applies for ψ if one assumes the contrary of Lemma 2, obtaining that

$$\frac{\psi(\alpha_i) - \psi(0)}{|\alpha_i|} \to -\infty, \quad i \to \infty. \tag{4.2}$$

But ψ is decreasing as a function of each component separately. Hence, ψ is differentiable a.e. and so (4.2) can only hold in a set of measure zero. By Lemma 2, Q_a is normal for almost all a in R^n.

5 TRANSLATIONS: THE LINEAR CASE

In this section we prove that, if the dynamics are linear with respect to x, most problems are normal with respect to translations of the end-point set. Let us start characterizing the notion of calmness.

4. **Lemma:** Suppose we are given a function $g:R^n \to R$ and a family of sets in R^n, $\{C_a \mid a \in R^n\}$. Let

$$\phi(a) := \inf\{g(x) \mid x \in C_a\}, \qquad Q := \{a \in R^n \mid C_a \neq \emptyset\}$$

and consider the following statements:

i. $$\liminf_{a \to 0} \frac{\phi(a) - \phi(0)}{|a|} > -\infty$$

ii. There exist K and $\delta > 0$ such that, for all $a \in Q \cap \delta B$,

$$\phi(a) - \phi(0) \geqslant -K|a|$$

where B denotes the open unit ball in R^n deleting zero.

Then, if $\phi(0) > -\infty$, (i) $\Rightarrow$ (ii) and, if $0 \in Q$, (ii) $\Rightarrow$ (i). In particular, if $\phi(0) \in R$ then (i) and (ii) are equivalent.

Proof: Define $f:R^n \to R$ as follows:

$$f(a) = \frac{\phi(a) - \phi(0)}{|a|} \quad \text{if } a \in R^n - \{0\}$$

$$= 0 \quad \text{if } a = 0$$

and let $Q_f := \{a \in R^n \mid f(a) < +\infty\}$. By definition we have that

$$\liminf_{a \to 0} f(a) > -\infty$$

if, and only if, there exist K and $\delta > 0$ such that, for all $a \in Q_f \cap \delta B$, $f(a) \geqslant -K$. The result follows observing that

$$Q = \{a \in R^n \mid \phi(a) < +\infty\}$$

for, if $\phi(0) > -\infty$, then $Q \subset Q_f$ and so (i) $\Rightarrow$ (ii); if $0 \in Q$, then $Q_f - \{0\} \subset Q$ and so (ii) $\Rightarrow$ (i).

Now we can express the notion of calmness in a very simple way. Set:

$A := \{x(1) \mid x \in X_*\}$

$Q := \{a \in R^n \mid A \cap (C + \{a\}) \neq \emptyset\}$

$S := \{a \in Q \mid$ there exist $K, \delta > 0$, such that, for all $b \in Q \cap (\delta B + \{a\})$

$$\phi(b) - \phi(a) \geq -K|b - a|\}.$$

Observe that, for all $a \in R^n$,

$$\phi(a) = \inf\{g(x) \mid x \in A \cap (C + \{a\})\}$$

and

$$Q = \{x - c \mid x \in A \text{ and } c \in C\} = A - C.$$

By Lemma 4 the notion of calmness is characterized by S as follows:

5. **Lemma:** P_a is calm if, and only if, $a \in S$ and $\phi(a) > -\infty$.

In other words, our problem is to see, roughly speaking, how "big" is S with respect to Q. Now, the set A is precisely the "attainable set" and, when the dynamics are linear with respect to x, this set is convex (in R^n). Before showing what happens in this case, let us recall a few concepts. A set $M \subset R^n$ is called an **affine set** or **linear variety** if

$$(1-\lambda)x + \lambda y \in M \text{ for every } x \in M,\ y \in M \text{ and } \lambda \in R,$$

the **affine hull** of a set $D \subset R^n$, denoted by aff D, is the intersection of all affine sets M such that $M \supset D$ and the **relative interior** of a convex set D (ri D), consists of the points x in aff D for which there exist $\varepsilon > 0$ such that $y \in D$ whenever $y \in \text{aff } D$ and $|x - y| \leqslant \varepsilon$.

The result due to I. Ekeland mentioned in Section 2 states that, if ϕ is lower semicontinuous on Q, then S is dense in Q (see [13]). It was observed by R. B. Vinter in [30] that, if A is compact, then the lower semicontinuity of ϕ on Q holds, $\phi(a)$ is finite for all a in Q and consequently, in view of Lemma 5, S is a dense set of points in Q where normality can be assured.

We present next a simple example of an abnormal problem. It illustrates some of the main difficulties that appear trying to impose conditions for calmness. It should be noted that the cost function is differentiable everywhere, the attainable set is convex and compact and the endpoint set is convex and closed.

Let $g(x,y) = x$ for all $(x,y) \in \mathbf{R} \times \mathbf{R}$, and consider the following sets:

$$A = \{(x,0) \mid -1 \leqslant x \leqslant 1\}$$

$$C = \{(x,y) \mid y \geqslant x^2 \}.$$

To show that normality is not satisfied, suppose $x^* = (x,y)$ solves a problem (P) where A is the attainable set and C the endpoint constraint. By the maximum principle, there exist an absolutely continuous function $p:[0,1] \to \mathbf{R} \times \mathbf{R}$ and $\lambda \in \{0,1\}$, not both zero, such that:

$$-p(1) \in N_C(x^*(1)) + \lambda \partial g(x^*(1)).$$

Now, observe that $x^*(1) = (x(1),y(1)) \in A \cap C$ and so $x(1) = y(1) = 0$. This implies that $N_C(x^*(1))$ is the half-line $x = 0$, $y \leqslant 0$, and, since $-p(1)$ is orthogonal to C at (0,0), it follows that $\lambda = 0$.

Since calmness implies normality, the problem posed is not calm. This can also be proved directly as follows: for all $0 \leqslant a \leqslant 1$, the value function at (0,-a) is given by:

$$\phi((0,-a)) = \inf\{g(x,y) \mid (x,y) \in A \cap (C + \{(0,-a)\})\}$$

$$= \inf\{x \in [-1,1] \mid x^2 \leqslant a\}$$

$$= - a^{1/2}.$$

So, the calmness condition fails, since:

$$\frac{\phi((0,-a)) - \phi((0,0))}{|(0,-a)|} = \frac{- a^{1/2}}{|a|} = \frac{-1}{a^{1/2}} \to -\infty, \quad a \to 0.$$

The result we prove below is that, if A and C are convex and A compact, then Q is convex and ri Q ⊂ S. If A is bounded but not closed, the same result holds assuming g is globally Lipschitz. We shall find it useful to first establish the following lemmas.

6. **Lemma:** Suppose A and C are convex, $0 \in \text{ri } Q$ (note that $Q = A - C$, and this is a convex set) and A is bounded. Then there exists $M > 0$ such that, for all $a \in Q$ and $x \in A \cap (C + \{a\})$, $\text{dist}(x , A \cap C) \leqslant M|a|$.

Proof: Since $0 \in \text{ri } Q$, there exists $\varepsilon > 0$ such that

$$b \in \text{aff } Q \text{ and } |b| \leqslant \varepsilon \Rightarrow b \in Q. \tag{4.3}$$

Since A is bounded, there exists R positive such that, for all x and y in A, $|x - y| \leqslant R$. Now, let $a \in Q$, and $x \in A \cap (C + \{a\})$. We shall show that the required inequality holds with $M = R/\varepsilon$.

Let $b := - \varepsilon \frac{a}{|a|}$ and note that, since $a \in Q \subset \text{aff } Q$ and $0 \in \text{aff } Q$, it follows that $b \in \text{aff } Q$. Also $|b| = \varepsilon$, and so, by (4.3), $b \in Q$. Consequently, $A \cap (C + \{b\}) \neq \emptyset$. Now, let $y \in A \cap (C + \{b\})$, c_0 and $c_1 \in C$ such that $x = c_0 + a$ and $y = c_1 + b$ and set $\lambda := |a|/(|a|+\varepsilon)$. Observe that:

$$c := \lambda c_1 + (1-\lambda)c_0 = \lambda y + (1-\lambda)x \quad \varepsilon \quad A \cap C.$$

The result now follows since

$$\text{dist}(x \ , \ A \cap C) \leqslant |c - x| = \lambda|x - y| \leqslant \frac{|a|}{\varepsilon} |x - y| \leqslant \frac{R}{\varepsilon} |a|.$$

7. **Lemma:** Suppose $0 \varepsilon Q$ and there exist M and $\delta > 0$ such that, for all $a \varepsilon Q \cap \delta B$ and all $x \varepsilon A \cap (C + \{a\})$,

$$\text{dist}(x \ , \ A \cap C) \leqslant M|a|.$$

If A is compact or g globally Lipschitz then $0 \varepsilon S$.

Proof: Let $a \varepsilon Q \cap \delta B$ and $\varepsilon > 0$. By definition of ϕ and the function distance, there exist $x \varepsilon A \cap (C + \{a\})$ such that $g(x) \leqslant \phi(a) + \varepsilon$ and $y \varepsilon A \cap C$ such that

$$|x - y| \leqslant \text{dist}(x, A \cap C) + |a| \leqslant (M+1)|a|.$$

Now, the assumptions A compact or g globally Lipschitz, imply the existence of some $L > 0$, independent of a and ε, such that

$$|g(x) - g(y)| \leqslant L|x - y|.$$

From this it follows that

$$\phi(a) - \phi(0) + \varepsilon \geqslant g(x) - g(y) \geqslant - L|x - y| \geqslant - L(M+1)|a|.$$

Since ε is arbitrary we obtain the required result.

We are now in a position to show that normality holds for all (linear) problems obtained by translating the original endpoint set, except possibly in directions that belong to the relative boundary of the set Q.

8. **Theorem:** Suppose A and C are convex and A bounded. If also A is closed or g is globally Lipschitz, then $\text{ri } Q \subset S$.

Proof: Let $a_0 \in \text{ri } Q$ and define, for all $b \in \mathbf{R}^n$,

$$\tilde{\phi}(b) := \inf\{g(x) \mid x \in A \cap ((C + \{a_0\}) + \{b\})\}.$$

Set $R := A - (C + \{a_0\})$ and observe that $a_0 \in \text{ri } Q \Rightarrow a_0 \in \text{ri } R + \{a_0\} \Rightarrow 0 \in \text{ri } R$. Consequently, by Lemmas 6 and 7 there exist K and $\delta > 0$ such that $b \in R \cap \delta B \Rightarrow \tilde{\phi}(b) - \tilde{\phi}(0) \geqslant -K|b|$. But $\phi(a_0 + b) = \tilde{\phi}(b)$ for all $b \in \mathbf{R}^n$. Hence, if $a \in Q \cap (\delta B + \{a_0\})$

$$b := a - a_0 \in R \cap \delta B$$

and

$$\phi(a) - \phi(a_0) = \tilde{\phi}(b) - \tilde{\phi}(0) \geqslant -K|a - a_0|.$$

This implies $a_0 \in S$ and the proof is complete.

CONCLUSIONS

1. Optimal Control Systems with Time Delay. For an optimal control problem with a delay in the phase coordinates, G. L. Kharatishvili (see [25]) established in 1961 necessary conditions in the form of a maximum principle. In subsequent works this principle was generalized for problems with several delays and the best known results were given in 1969 by S. C. Huang ([19]). These conditions imply straightforward the necessary conditions of Euler, Legendre and Weierstrass as stated in Chapter 2. An equivalent of Jacobi's necessary condition was not known and sufficiency was neglected. The reason, as pointed out by G. Tadmor in [29] and explained in detail in Section 2.5, is that the Jacobi equation becomes a differential-difference equation involving both advanced and retarded arguments.

The first contribution of this research is the content of Theorems 2.3 and 2.5: necessary and sufficient conditions expressed explicitly in terms of the variations. There are problems where the second order condition for optimality is readily verified, such as example 5.2, and this results become applicable.

For most problems, it is necessary to have a verifiable characterization of the positivity of the second variation. This is usually accomplished by a direct computation of the fundamental matrix solution of the Jacobi equation, through fields of extremals or solving a matrix

Riccati equation. We centered our attention in the first procedure. The main result of this section is that there do exist initial conditions that guarantee existence and uniqueness of the solution of Jacobi's equation. This is proved in Theorem 2.13, where an extra assumption is required: the delay should be "sufficiently small". A conjecture at this point is clear: to see if this assumption is essential, i.e., if the same result holds if the delay is not assumed to be bounded. Now, even if this fact were true, we would still have problems verifying conditions (C1)-(C3) of the sufficient conditions stated in Theorem 2.22, since they are based on a nonsingular matrix solution of the Jacobi equation obtained only implicitly throughout the proof.

The method of steps in Chapter 3 avoids these difficulties and sufficiency is verifiable solving a matrix Riccati inequality or Riccati equation. The open question in this approach consists in finding a corresponding necessary condition to the one we called "Riccati condition". One suspects that it exists, since the necessary condition "there are no conjugate sequences on (t_0,t_1)" exploits the nonnegativity of $I''(x_0;\cdot)$ only along $\hat{Y}$, without taking into account trajectories nonvanishing at the endpoints 0 and θ.

All these results could be extended to optimal control systems involving several delays in a nonsmooth analysis context.

2. **Normality Conditions.** Motivated by the results obtained for a mathematical programming problem ("most problems are normal"), it sounds acceptable the conjecture that most problems will be normal in optimal control. This question has remained unanswered since Pontryagin's principle was established. We prove that indeed this is the case if one perturbs the endpoint set by enlarging it. For problems that are obtained by translating this set, the question was reduced to the following: suppose we are given a locally Lipschitz function $g:R^n \to R$ and two sets A

and C in R^n. Define for all $a \in R^n$:

$$\phi(a) := \inf \{g(x) \mid x \in A \cap (C + \{a\})\}$$

$$Q := \{a \in R^n \mid A \cap (C + \{a\}) \neq \emptyset\}$$

and let S be the set of points $a \in Q$ for which

$$\liminf_{b \to a} \frac{\phi(b) - \phi(a)}{|b - a|} > -\infty.$$

The question is reduced to find how "big" is S with respect to Q. In other words, when is ϕ lower semi-Lipschitz continuous. Given a problem for which A (the attainable set) is compact (this is satisfied by the so-called "relaxed problems"), it was shown by R. B. Vinter ([30]) using a result of I. Ekeland ([13]), that S (the set of directions where normality can be guaranteed) is dense in Q (the set of directions for which the attainable set translated and the endpoint constraint set intersect). We extend this result when the dynamics are linear in the state variable, proving that Q is convex, and the relative interior of Q is contained in S. In other words, all relaxed and linear problems obtained by translating the endpoint set are normal, except possibly in directions that belong to the relative boundary of Q. The non-linear case remains unanswered but future work studying properties of the value function could solve the question of normality under translations.

BIBLIOGRAPHY

[1] J. P. AUBIN and I. EKELAND: "Applied Nonlinear Analysis", John Wiley & Sons, New York, 1984.

[2] R. BELLMAN and K. L. COOKE: "Differential-Difference Equations", Academic Press, New York, 1963.

[3] L. D. BERKOVITZ: "Optimal Control Theory", Springer-Verlag, New York, 1974.

[4] G. A. BLISS: "Lectures on the Calculus of Variations", University of Chicago Press, Chicago, 1946.

[5] O. BOLZA: "Lectures on the Calculus of Variations", Dover Publications, New York, 1961.

[6] L. CESARI: "Optimization- Theory and Applications", Springer-Verlag, New York, 1983.

[7] F. H. CLARKE: "A New Approach to Lagrange Multipliers", Math. Oper. Res., Vol. 1, No. 2, (1976) 165-174.

[8] -----: "The Generalized Problem of Bolza", SIAM J. Control Optim., Vol. 14, (1976) 682-699.

[9] -----: "The Maximum Principle under Minimal Hypotheses", SIAM J. Control Optim., Vol. 14, (1976) 1078-1091.

[10] -----: "Optimization and Nonsmooth Analysis", John Wiley & Sons, New York, 1983.

[11] F. H. CLARKE and V. ZEIDAN: "Sufficiency and the Jacobi Condition in the Calculus of Variations", (to appear), 1985.

[12] R. D. DRIVER: "Ordinary and Delay Differential Equations", Springer-Verlag, New York, 1977.

[13] I. EKELAND: "Nonconvex Minimization Problems", Bull. Am. Math. Soc., Vol. 1, No. 3, (1979) 443-474.

[14] I. EKELAND and G. LEBOURG: "Generic Fréchet Differentiability and Perturbed Optimization Problems", Trans. Amer. Math. Soc., Vol. 224, (1976) 193-215.

[15] W. H. FLEMING and R. W. RISHEL: "Deterministic and Stochastic Optimal Control", Springer-Verlag, New York, 1975.

[16] H. HALKIN: "On the Necessary Condition for Optimal Control of Non-linear Systems", J. Analyse Math., Vol. 12, (1964) 1-82.

[17] H. HERMES and J. P. LASALLE: "Functional Analysis and Time Optimal Control", Academic Press, New York, 1969.

[18] M. R. HESTENES: "Calculus of Variations and Optimal Control Theory", John Wiley & Sons, New York, 1966.

[19] S. C. HUANG: "Optimal Control Problems with Retardations and Restricted Phase Coordinates", JOTA, Vol. 3, No. 5, (1969) 316-360.

[20] A. D. IOFFE and V. M. TIHOMIROV: "Theory of Extremal Problems", North-Holland Publishing Company, The Netherlands, 1979.

[21] E. B. LEE and L. MARKUS: "Foundations of Optimal Control Theory", John Wiley & Sons, New York, 1967.

[22] G. LEITMANN: "The Calculus of Variations and Optimal Control", Plenum Press, New York, 1981.

[23] D. G. LUENBERGER: "Optimization by Vector Space Methods", John Wiley & Sons, New York, 1969.

[24] M. MORSE: "Global Variational Analysis: Weierstrass Integrals on a Riemannian Manifold", Princeton University Press, Princeton, New Jersey, 1976.

[25] L. S. PONTRYAGIN et al.: "The Mathematical Theory of Optimal Processes", Wiley Interscience, New York, 1962.

[26] M. REED and B. SIMON: "Functional Analysis", Academic Press, New York, 1972.

[27] R. T. ROCKAFELLAR: "Convex Analysis", Princeton University Press, Princeton, New Jersey, 1970.

[28] J. E. SPINGARN and R. T. ROCKAFELLAR: "The Generic Nature of Optimality Conditions in Nonlinear Programming", Math. Oper. Res., Vol. 4, No. 4, (1979) 425-430.

[29] G. TADMOR: "The Structure of Hereditary and Nonlocal Systems", PhD Thesis, Weizmann Institute, Rehovot, Israel, 1984.

[30] R. B. VINTER: "The Equivalence of 'Strong Calmness' and 'Calmness' in Optimal Control Theory", J. Math. Anal. Appl., Vol. 96, No. 1, (1983) 153-179.

[31] -----: "New Global Optimality Conditions in Optimal Control Theory", SIAM J. Control Optim., Vol. 21, No. 2, (1983) 235-245.

[32] J. WARGA: "Optimal Control of Differential and Functional Equations", Academic Press, New York, 1972.

[33] L. C. YOUNG: "Lectures on the Calculus of Variations and Optimal Control Theory", Saunders, Philadelphia, 1969.