HUGHES, David Knox, 1940-
CONTRIBUTIONS TO THE THEORY OF VARIATIONAL AND OPTIMAL CONTROL PROBLEMS WITH DELAYED ARGUMENT.

The University of Oklahoma, Ph.D., 1967
Mathematics

THE UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

CONTRIBUTIONS TO THE THEORY OF VARIATIONAL AND OPTIMAL CONTROL PROBLEMS WITH DELAYED ARGUMENT

A DISSERTATION

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

degree of

DOCTOR OF PHILOSOPHY

BY

DAVID KNOX HUGHES

Norman, Oklahoma

1967

CONTRIBUTIONS TO THE THEORY OF VARIATIONAL AND OPTIMAL CONTROL PROBLEMS WITH DELAYED ARGUMENT

APPROVED BY

George M. Ewing

W T Reid

R S Davis

John E. Driver

Gene Levy

DISSERTATION COMMITTEE

ACKNOWLEDGMENT

The author gratefully expresses his appreciation for the invaluable advice and guidance of Professor George M. Ewing who directed the writing of this paper. Thanks is also due Professor W. T. Reid for many helpful comments and suggestions.

A portion of the research reported here was done while the author was a Research Assistant at the University of Oklahoma Research Institute working under grants AF-AFOSR-211-63 and AF-AFOSR-749-65 from the Air Force Office of Scientific Research.

TABLE OF CONTENTS

CONTRIBUTIONS TO THE THEORY OF VARIATIONAL AND OPTIMAL CONTROL PROBLEMS WITH DELAYED ARGUMENT

CHAPTER I

INTRODUCTION

This paper is devoted to discovering necessary conditions together with some sufficient conditions for optimality in variational and control problems with delayed argument, that is, problems which involve functionals of the type found in relation (1.1). In recent years there have been many articles and several books which deal with various aspects of such problems. In particular we call attention to the books of M. N. Oğuztöreli [10] and of L. E. El'sgol'c (also transliterated Elsgolts) [3]. Oğuztöreli discusses delay-differential equations in some detail and then studies the question of existence of optimal controls for delay-differential control systems. He also discusses necessary conditions from the point of view of dynamic programming. El'sgol'c gives a few necessary conditions for a minimum problem of a more classical nature [3, p.215] than that of Oğuztöreli [10, p.171]. We study neither the problem of Oğuztöreli nor the problem of El'sgol'c although our problems do have similarities to both. There is no one formulation among those that the present author has encountered which can be identified as the canonical or standard problem involving delays. The problems that receive most attention in this paper are those among various others examined by the author for which he has been able to obtain a collection of results comparable to corresponding parts of the theory of necessary conditions and of sufficient conditions for classical problems of the calculus of variations with no delays. Much of the published work on problems with delay has thus far been in existence theory. There are no published results insofar as the author is aware on necessary conditions analogous to those of Weierstrass and Jacobi for classical problems. Neither has any sufficient condition for local or global extrema been given except that in [2, p.556] which appeared subsequently to most of the work reported here.

Although differential equations with delays have been investigated in occasional papers over a number of years, the wide recent and current interest in general systems theory and in optimal design and control of electromechanical systems in weaponry and industry together with problems in mathematical economics and in other areas has motivated the introduction of variational problems with delays and the expanded recent literature on delay-differential equations.

The objective of the second and third chapters of this paper is an investigation of a functional $J(y)$ such that

$$(1.1) \qquad J(y) = \int_a^b f[t, y(t-\tau), y(t), \dot{y}(t-\tau), \dot{y}(t)]dt.$$

There are no side-conditions; τ is a positive real number; and y is a continuous piecewise smooth vector function with n components. We find necessary conditions analogous to those of Euler, Weierstrass, and Legendre for the classical fixed endpoint problem [1]. Also a fourth necessary condition involving proper values associated with a certain boundary value problem is derived. A sufficient condition patterned after that of Ewing [4] is obtained, and the indirect method of Hestenes is used to obtain sufficiency in a special case.

The fourth chapter contains a maximum principle for a problem with time lag similar to the control problem without lag considered by Hestenes [7] and also similar to the problem discussed in the important book by L. S. Pontryagin et. al. [11, p.213]. The approach used is that of Hestenes. In particular we investigate the functional $I_0(y)$ where

$$I_0(y) = \int_a^b L_0[t,y(t-\tau),y(t),u(t)]dt \tag{1.2}$$

is to be minimized on a class of functions satisfying the conditions

$$\dot{y}^i = f^i[t,y(t-\tau),y(t),u(t)], \quad i = 1,\ldots,n, \quad a \le t \le b;$$

$$y^i(t) = \alpha^i(t), \quad a-\tau \le t \le a; \quad y^i(b) = \text{constant}, \quad i = 1,\ldots,n.$$

The vector function $y = (y^1,\ldots,y^n)$ is also subject to the isoperimetric conditions

$$I_\gamma(y) = \int_a^b L[t,y(t-\tau),y(t),u(t)]dt \le 0, \quad 1 \le \gamma \le p',$$

$$I_\gamma(y) = \int_a^b L[t,y(t-\tau),y(t),u(t)]dt = 0, \quad p' < \gamma \le p.$$

We also give an indirect sufficiency proof for a slight modification of the above problem.

In the remainder of this paper we use the abbreviations PWS for piecewise smooth and PWC for piecewise continuous. By a PWS function on [a,b] we mean a continuous function which has PWC derivatives on [a,b]. This class of functions is sometimes denoted by the symbol D'[a,b]. We adopt the convention that a repeated index will specify summation on the index unless specifically stated otherwise.

Let Ω be a suitable class of PWS vector functions defined on an interval [a,b]. For $J : \Omega \to R$, we wish to define minima of J on Ω. First let x and y be elements of Ω. Define a strong distance ρ_s and a weak distance ρ_w as follows:

$$\rho_s(x,y) \equiv \sup |x(t) - y(t)|, \; t \text{ in } [a,b];$$

$$\rho_w(x,y) \equiv \sup |\dot{x}(t) - \dot{y}(t)| + \rho_s(x,y), \; t \text{ in } [a,b]^*$$

where $| \; |$ denotes the Euclidean norm and

$$[a,b]^* = \{t : t \text{ in } [a,b]; \; \dot{y}(t), \dot{x}(t) \text{ exist}\}.$$

We now say that J(y) has a <u>weak local minimum on</u> Ω <u>at</u> y_0 if and only if there exists $\delta > 0$ such that

(1.3) $J(y_0) \leq J(y)$ for all y in Ω such that $0 < \rho_w(y,y_0) < \delta$.

We say that $J(y_0)$ is a <u>strong local minimum on</u> Ω if and only if there exists $\delta > 0$ such that

(1.4) $J(y_0) \leq J(y)$ for all y in Ω such that $0 < \rho_s(y,y_0) < \delta$.

$J(y_0)$ is a <u>global minimum on</u> Ω if and only if

(1.5) $$J(y_0) \leq J(y) \text{ for all } y \text{ in } \Omega.$$

Clearly if y_0 furnishes a minimum for $J(y)$ in the sense of (1.5), then it furnishes a minimum in the sense of (1.4) and hence in the sense of (1.3). By the phrase "y_0 minimizes $J(y)$" we will mean minimization in one of the senses (1.3), (1.4), (1.5) and hence in the sense of (1.3).

CHAPTER II

NECESSARY CONDITIONS FOR THE SIMPLE INTEGRAL PROBLEM

2.1 Introduction

It is the purpose of this chapter to find necessary conditions on a function y which minimizes the functional

$$(2.1) \qquad J(y) = \int_a^b f\big(t, y(t - \tau), y(t), \mathring{y}(t - \tau), \mathring{y}(t)\big)dt$$

on the class Ω of all PWS vector functions $y = (y^1, \ldots, y^n)$ such that $y^i(t) = \alpha^i(t)$, $a - \tau \leq t \leq a$, $i = 1, \ldots, n$, where $\alpha(t) = \big(\alpha^i(t)\big)$ is a given PWS vector function and also such that $y^i(b) = \beta^i$ = constant, $i = 1, \ldots, n$. We assume that $f(t,x,y,q,r)$ is continuous on the region $\mathcal{R}: [a,b] \times R^n \times R^n \times R^n \times R^n$ and has continuous partial derivatives of the first two orders with respect to the variables (x,y,q,r). The constant τ is positive and $\tau < b - a$. If $\tau \geq b - a$, then since $y(t)$ is fixed on $a - \tau \leq t \leq a$, the problem reduces to the classical fixed endpoint problem. We define the symbols x and z by the formulas $x(t) = y(t - \tau)$, $z(t) = y(t + \tau)$. The following convention is in effect throughout this entire chapter.

CONVENTION 2.1.1. <u>When a condition involves the interval</u> $[a,b]$ <u>and any of the symbols</u> $\mathring{x}(t)$, $\mathring{y}(t)$ <u>or</u> $\mathring{z}(t)$, <u>then at any interior point of</u> $[a,b]$ <u>where one or more of these derivatives fail to exist, the stated condition is understood to hold with the derivatives interpreted</u>

as either right or left derivatives.

It should be noted that most of what is done in this chapter and in Chapter 3 remains valid if τ is replaced by a function $\tau(t)$ with suitable restrictions. We choose to consider the case with τ a constant since the notation is much simpler.

2.2 An Euler Equation

THEOREM 2.2.1 *If y in Ω furnishes J with a minimum, then there exist constants c_i, $i = 1,\ldots,n$, such that y must satisfy the following integro-differential-difference equations:*

$$(2.2a)\qquad f_{r^i}(t,x,y,\mathring{x},\mathring{y}) + f_{q^i}(t+\tau,y,z,\mathring{y},\mathring{z}) = \int_{b-\tau}^{t}[f_{y^i}(s,x,y,\mathring{x},\mathring{y}) + f_{x^i}(s+\tau,y,z,\mathring{y},\mathring{z})]ds + c_i, \quad a \le t \le b-\tau;$$

$$(2.2b)\qquad f_{r^i}(t,x,y,\mathring{x},\mathring{y}) = \int_{b-\tau}^{t} f_{y^i}(s,x,y,\mathring{x},\mathring{y})ds + c_i, \quad b-\tau \le t \le b,$$

in which $x,y,z,\mathring{x},\mathring{y},\mathring{z}$ are respective abbreviations for $x(t)$ or $x(s)$ etc.

In order to prove this result let η be a PWS vector function defined on $a-\tau \le t \le b$ satisfying the conditions $\eta(t) \equiv 0$, $a-\tau \le t \le a$, $\eta(b) = 0$. Consider the function

$$F(\varepsilon) = J(y+\varepsilon\eta) = \int_a^b f(t,x+\varepsilon\xi,y+\varepsilon\eta,\mathring{x}+\varepsilon\mathring{\xi},\mathring{y}+\varepsilon\mathring{\eta})dt,$$

where $\xi(t) = \eta(t-\tau)$. Taking the derivative of $F(\varepsilon)$ at $\varepsilon = 0$, we find that

$$(2.3)\qquad F'(0) = \int_a^b (f_{x^i}\xi^i + f_{y^i}\eta^i + f_{q^i}\mathring{\xi}^i + f_{r^i}\mathring{\eta}^i)dt = 0$$

where the arguments of the integrand functions are $(t, x(t), y(t), \mathring{x}(t), \mathring{y}(t))$.

By a linear change of variable in the first and third terms under the integral, (2.3) becomes

$$\int_a^b [\phi_i(t)\eta^i(t) + \psi_i(t)\mathring{\eta}^i(t)]dt = 0 \tag{2.4}$$

where

$$\phi_i(t) = f_{y^i}(t,x,y,\mathring{x},\mathring{y}) + f_{x^i}(t+\tau,y,z,\mathring{y},\mathring{z}),\ a \le t \le b-\tau,$$

$$\phi_i(t) = f_{y^i}(t,x,y,\mathring{x},\mathring{y}),\quad b-\tau < t \le b,$$

and

$$\psi_i(t) = f_{r^i}(t,x,y,\mathring{x},\mathring{y}) + f_{q^i}(t+\tau,y,z,\mathring{y},\mathring{z}),\ a \le t \le b-\tau,$$

$$\psi_i(t) = f_{r^i}(t,x,y,\mathring{x},\mathring{y}),\quad b-\tau < t \le b.$$

Integration of (2.4) by parts and use of the boundary conditions on η yields the equations

$$\int_a^b [\psi_i(t) - \int_{b-\tau}^t \phi_i(s)ds]\mathring{\eta}^i(t)dt = 0,\ i = 1,\ldots,n. \tag{2.5}$$

Applying the du Bois-Reymond Lemma [1, p.10] to (2.5) one has the analog of the classical Euler equations:

$$\psi_i(t) = \int_{b-\tau}^t \phi_i(s)ds + c_i,\ i = 1,\ldots,n;\ a \le t \le b. \tag{2.6}$$

After substitution for ϕ_i and ψ_i this yields the stated equations (2.2).

COROLLARY 2.2.1 <u>If</u> y <u>in</u> Ω <u>furnishes</u> J <u>with a minimum, then</u> y <u>must satisfy the following differential-difference equations for all values of</u> t <u>on</u> $[a,b]$ <u>except the possible finite set of t-values which correspond to corners of</u> x, y, <u>or</u> z:

(2.7a) $$f_{y^i}(t,x,y,\mathring{x},\mathring{y}) + f_{x^i}(t+\tau,y,z,\mathring{y},\mathring{z}) =$$

$$\frac{d}{dt}[f_{r^i}(t,x,y,\mathring{x},\mathring{y}) + f_{q^i}(t+\tau,y,z,\mathring{y},\mathring{z})],\ i = 1,\ldots,n;\ a \le t \le b - \tau;$$

(2.7b) $$f_{y^i}(t,x,y,\dot{x},\dot{y}) = \frac{d}{dt} f_{r^i}(t,x,y,\dot{x},\dot{y}),\ i = 1,\ldots,n;\ b - \tau \le t \le b.$$

To prove this result, differentiate (2.2).

COROLLARY 2.2.2 <u>If</u> y <u>in</u> Ω <u>furnishes</u> J <u>with a minimum, then at</u> $t = b - \tau$ <u>the following relation holds</u>:

(2.8) $$f_{r^i}[b-\tau,x(b-\tau),y(b-\tau),\mathring{x}_-(b-\tau),\mathring{y}_-(b-\tau)] +$$

$$f_{q^i}[b,y(b-\tau),z(b-\tau),\mathring{y}_-(b-\tau),\mathring{z}_-(b-\tau)] =$$

$$f_{r^i}[b-\tau,x(b-\tau),y(b-\tau),\mathring{x}_+(b-\tau),\mathring{y}_+(b-\tau)],\ i = 1,\ldots,n,$$

<u>in which</u> $\dot{x}_-$, $\mathring{x}_+$ <u>etc. denote respective left and right derivatives.</u>

The stated conclusion is immediate from relations (2.6), (2.2) and the continuity of the integral in t.

Also from relations (2.2) one obtains the further (Erdmann) corner conditions if t is the abscissa of a corner:

$$(2.9a)\quad f_{r^i}[t,x(t),y(t),\dot{x}_-(t),\dot{y}_-(t)] + f_{q^i}[t+\tau,y(t),z(t),\dot{y}_-(t),\dot{z}_-(t)] =$$

$$f_{r^i}[t,x(t),y(t),\dot{x}_+(t),\dot{y}_+(t)] + f_{q^i}[t+\tau,y(t),z(t),\dot{y}_+(t),\dot{z}_+(t)],\quad a < t < b-\tau;$$

$$(2.9b)\quad f_{r^i}[t,x(t),y(t),\dot{x}_-(t),\dot{y}_-(t)] = f_{r^i}[t,x(t),y(t),\dot{x}_+(t),\dot{y}_+(t)],\quad b-\tau < t < b.$$

We now consider several special cases. If

$$(2.10)\qquad f(t,x,y,q,r) = g(t,x,y,r) + h(t,x,y,q),$$

then the Euler equations (2.2) become

$$(2.11a)\qquad g_{r^i} + h_{q^i} = \int_{b-\tau}^{t}(g_{y^i} + h_{y^i} + g_{x^i} + h_{x^i})ds + c_i,\quad i = 1,\ldots,n;\ a \le t \le b-\tau;$$

$$(2.11b)\quad g_{r^i} = \int_{b-\tau}^{t}(g_{y^i} + h_{y^i})ds + c_i,\quad i = 1,\ldots,n;\ b-\tau \le t \le b.$$

The partial derivatives g_{r^i}, g_{y^i}, are evaluated at $(t,x(t),y(t),\dot{y}(t))$; h_{y^i} stands for $h_{y^i}(t,x(t),y(t),\dot{x}(t))$; the partial derivatives g_{x^i} are evaluated at $(t+\tau,y(t),z(t),\dot{z}(t))$; and the partial derivativex h_{x^i} and h_{q^i} are evaluated at $(t+\tau,y(t),z(t),\dot{y}(t))$. In this case the corner condition (2.8) at $t = b-\tau$ becomes

$$(2.12)\quad g_{r^i}[b-\tau,x(b-\tau),y(b-\tau),\dot{y}_-(b-\tau)] + h_{q^i}[b,y(b-\tau),z(b-\tau)\dot{y}_-(b-\tau)] =$$

$$g_{r^i}[b-\tau,x(b-\tau),y(b-\tau),\dot{y}_+(b-\tau)],\quad i = 1,\ldots,n.$$

If t is the abscissa of a corner of y for $a < t < b - \tau$, then corresponding to (2.9a) we have the relation

$$(2.13a)\quad g_{r^i}[t,x(t),y(t),\dot{y}_-(t)] + h_{q^i}[t+\tau,y(t),z(t),\dot{y}_-(t)] =$$

$$g_{r^i}[t,x(t),y(t),\dot{y}_+(t)] + h_{q^i}[t+\tau,y(t),z(t),\dot{y}_+(t)],\quad i = 1,\ldots,n.$$

If $b - \tau < t < b$, then corresponding to (2.9b) we have

$$(2.13b)\quad g_{r^i}[t,x(t),y(t),\dot{y}_-(t)] = g_{r^i}[t,x(t),y(t),\dot{y}_+(t)],\quad i = 1,\ldots,n.$$

Now let $f(t,x,y,q,r) = m(t,y,r) + n(t,x,q)$. We will call this special case the <u>separated</u> problem. In this case the Euler equations have no delayed arguments:

$$(2.14a)\quad m_{r^i}(t,y,\dot{y}) + n_{q^i}(t+\tau,y,\dot{y}) = \int_{b-\tau}^{t}[m_{y^i}(s,y,\dot{y}) + n_{x^i}(s+\tau,y,\dot{y})]ds + c_i,$$
$$a \le t \le b - \tau;$$

$$(2.14b)\quad m_{r^i}(t,y,\dot{y}) = \int_{b-\tau}^{t} m_{y^i}(s,y,\dot{y})ds + c_i,\quad b - \tau \le t \le b,\ i = 1,\ldots,n.$$

There is a condition similar to (2.12) at $t = b - \tau$. Also of course there are corner conditions of the nature of (2.13a) and (2.13b). The separated problem will be quite useful in constructing some examples.

2.3 <u>Legendre Condition</u>

Again we consider the problem described in Section 2.1, and we assume that y is a minimizing function for $J(y)$. Since $\varepsilon = 0$ furnishes $F(\varepsilon)$ with a minimum, it necessarily follows that $F''(0) \ge 0$. Hence

$$(2.15)\quad \int_a^b [f_{x^ix^j}\xi^i\xi^j + 2f_{x^iy^j}\xi^i\eta^j + 2f_{x^iq^j}\xi^i\dot\xi^j + 2f_{x^ir^j}\xi^i\dot\eta^j + f_{y^iy^j}\eta^i\eta^j +$$

$$2f_{y^iq^j}\eta^i\dot\xi^j + 2f_{y^ir^j}\eta^i\dot\eta^j + f_{q^iq^j}\dot\xi^i\dot\xi^j + 2f_{q^ir^j}\dot\xi^i\dot\eta^j + f_{r^ir^j}\dot\eta^i\dot\eta^j]dt \geq 0.$$

The arguments of the functions in (2.15) are $(t, x(t), y(t), \dot x(t), \dot y(t))$.

Let t_0 be an arbitrary point in $(a, b - \tau)$ such that neither $t_0 - \tau$, t_0, nor $t_0 + \tau$ corresponds to a corner of y. Choose $\delta > 0$ such that

$$\delta < \min\left\{\frac{\tau}{2}, t_0 - a, b - \tau - t_0\right\}$$

and such that there is no value of t corresponding to a corner of y in the intervals $(t_0 - \tau - \delta, t_0 - \tau + \delta)$, $(t_0 - \delta, t_0 + \delta)$, $(t_0 + \tau - \delta, t_0 + \tau + \delta)$. Let $\pi = (\pi^i)$ be a constant vector and define the vector function $\eta(t) = \left(\eta^i(t)\right)$ by the formulas

$$\eta^i(t) = \begin{cases} 0, & t \notin (t_0 - \delta, t_0 + \delta) \\ \delta\pi^i\left[1 - \dfrac{|t_0 - t|}{\delta}\right], & t \in (t_0 - \delta, t_0 + \delta),\ i = 1, \ldots, n. \end{cases}$$

This is an admissible η, and by use of it (2.15) becomes

$$(2.16)\quad \int_{t_0-\delta}^{t_0+\delta} (f_{y^iy^j}\eta^i\eta^j + 2f_{x^iy^j}\xi^i\eta^j + 2f_{y^ir^j}\eta^i\dot\eta^j + 2f_{y^iq^j}\eta^i\dot\xi^j$$

$$+ 2f_{x^ir^j}\xi^i\dot\eta^j + f_{r^ir^j}\dot\eta^i\dot\eta^j + 2f_{r^iq^j}\dot\eta^i\dot\xi^j)dt$$

$$+ \int_{t_0+\tau-\delta}^{t_0+\tau+\delta} (f_{x^ix^j}\xi^i\xi^j + 2f_{x^iq^j}\xi^i\dot\xi^j + f_{q^iq^j}\dot\xi^i\dot\xi^j)dt$$

$$=\int_{t_0-\delta}^{t_0+\delta}\{[f_{y^iy^j}(\ ) + f_{x^ix^j}(+\tau)]\eta^i\eta^j + 2[f_{y^i\dot{r}^j}(\ ) + f_{x^iq^j}(+\tau)]\eta^i\dot{\eta}^j$$

$$+ [f_{r^ir^j}(\ ) + f_{q^iq^j}(+\tau)]\dot{\eta}^i\dot{\eta}^j\}dt \geq 0,$$

in which $(\ )$ *stands for* $[t,x(t),y(t),\dot{x}(t),\dot{y}(t)]$ and $(+\tau)$ stands for $[t+\tau,y(t),z(t),\dot{y}(t),\dot{z}(t)]$. The terms in $\eta^i\xi^j$, $\dot{\eta}^i\xi^j$, $\eta^i\dot{\xi}^j$, $\dot{\eta}^i\dot{\xi}^j$ disappear since $\xi^i(t) \equiv 0$ for $t_0 - \delta < t < t_0 + \delta$, $i = 1,\ldots,n$.

Substituting for η^i in (2.16), we have that

$$(2.17)\quad \int_{t_0-\delta}^{t_0+\delta}\{P_{ij}(t)\pi^i\pi^j\ \delta^2\left[1 - \frac{|t - t_0|}{\delta}\right]^2$$

$$\pm\ 2Q_{ij}(t)\pi^i\pi^j\ \delta\left[1 - \frac{|t - t_0|}{\delta}\right] + (\pm 1)^2R_{ij}(t)\pi^i\pi^j\}dt \geq 0$$

where $P_{ij}(t)$, $Q_{ij}(t)$, $R_{ij}(t)$ stand for the bracketed factors inthe preceeding integral. Dividing (2.17) by 2δ and taking the limit as δ tends to zero, we conclude that $R_{ij}(t_0)\pi^i\pi^j \geq 0$. If t_0, $t_0 - \delta$, or $t_0 + \delta$ corresponds to a corner of y or if $t_0 = a$ or $t_0 = b - \tau$, the inequality

$$R_{ij}(t_0)\pi^i\pi^j \geq 0$$

still holds in the sense of convention 2.1.1 since $R(t)$ is continuous.

If t_0 is in the interval $(b - \tau,b)$ such that neither t_0 nor $t_0 - \tau$ corresponds to a corner of y, then $F''(0)$ becomes (with η defined as before)

$$\int_{t_0-\delta}^{t_0+\delta} (f_{y^iy^j}\eta^i\eta^j + 2f_{x^iy^j}\xi^i\eta^j + 2f_{y^ir^j}\eta^i\mathring{\eta}^j + 2f_{y^iq^j}\eta^i\mathring{\xi}^j$$

$$+ 2f_{x^ir^j}\xi^i\mathring{\eta}^j + f_{r^ir^j}\mathring{\eta}^i\mathring{\eta}^j + 2f_{r^iq^j}\mathring{\eta}^i\mathring{\xi}^j)dt \geq 0.$$

Setting $R_{\mathring{i}\mathring{j}}(t) = f_{r^ir^j}(t,x(t),y(t),\mathring{x}(t),\mathring{y}(t))$, one finds by the same analysis as before that $R_{\mathring{i}\mathring{j}}(t_0)\pi^i\pi^j \geq 0$. Again if t_0 or $t_0 - \tau$ corresponds to a corner of y or if $t_0 = b - \tau$ or $t_0 = b$, the inequality $R_{\mathring{i}\mathring{j}}(t\ )\pi^i\pi^j \geq 0$ remains true when interpreted in the sense of Convention 2.1.1.

Hence we have proved the following:

THEOREM 2.3.1 <u>If</u> y <u>furnishes</u> J <u>with a minimum, then it is necessary that along</u> y <u>the relation</u>

(2.18) $$R_{\mathring{i}\mathring{j}}(t)\pi^i\pi^j \geq 0, \quad a \leq t \leq b$$

<u>hold for every constant vector</u> π .

Here

(2.19a) $$R_{\mathring{i}\mathring{j}}(t) = f_{r^ir^j}(t,x(t),y(t),\mathring{x}(t),\mathring{y}(t))$$

$$+ f_{q^iq^j}(t+\tau,y(t),z(t),\mathring{y}(t),\mathring{z}(t)), \quad a \leq t \leq b - \tau;$$

(2.19b) $$R_{\mathring{i}\mathring{j}}(t) = f_{r^ir^j}(t,x(t),y(t),\mathring{x}(t),\mathring{y}(t)), \quad b - \tau < t \leq b.$$

If the function f is of the special form (2.10), the condition (2.18) becomes

(2.20a) $[g_{\dot{r}^i\dot{r}^j}(t,x,y,\dot{y}) + h_{\dot{q}^i\dot{q}^j}(t+\tau,y,z,\dot{y})]\pi^i\pi^j \geq 0, \quad a \leq t \leq b - \tau;$

(2.20b) $g_{\dot{r}^i\dot{r}^j}(t,x,y,\dot{y})\pi^i\pi^j \geq 0, \quad b - \tau \leq t \leq b.$

THEOREM 2.3.2. Suppose that f is of the form (2.10) and that y satisfies the Euler equations (2.11). Suppose also that for all p in R^n and $\pi \neq 0$

$$[g_{\dot{r}^i\dot{r}^j}(t,x,y,p) + h_{\dot{q}^i\dot{q}^j}(t+\tau,y,z,p)]\pi^i\pi^j > 0, \quad a \leq t \leq b - \tau.$$

Then y does not have a corner on $(a,b - \tau)$. Similarly if for all p in R^n and $\pi \neq 0$

$$g_{\dot{r}^i\dot{r}^j}(t,x,y,p)\pi^i\pi^j > 0, \quad b - \tau \leq t \leq b$$

then y does not have a corner on $(b - \tau,b)$.

Suppose that there exists t in $(a,b - \tau)$ such that $\dot{y}_-(t) \neq \dot{y}_+(t)$. Let $\dot{y}^i_-(t) = u^i$, $\dot{y}^i_+(t) = v^i$. Utilizing corner condition (2.13a), we have that

$$g_{\dot{r}^i}(t,x,y,u) - g_{\dot{r}^i}(t,x,y,v) + h_{\dot{q}^i}(t+\tau,y,z,u) - h_{\dot{q}^i}(t+\tau,y,z,v) = 0.$$

After an application of the mean value theorem this becomes

$$[g_{\dot{r}^i\dot{r}^j}(t,x,y,u+\theta(v-u)) + h_{\dot{q}^i\dot{q}^j}(t+\tau,y,z,u+\theta(v-u))](v^i-u^i)(v^j-u^j) = 0$$
$$0 < \theta < 1.$$

This contradicts the hypothesis since the vector $u-v$ is not the zero vector.

Similarly if $b-\tau < t < b$, we arrive at the expression

$$g_{\dot{r}^i\dot{r}^j}(t,x,y,u+\theta(v-u))(v^i-u^i)(v^j-u^j) = 0, \quad 0 < \theta < 1.$$

This is also a contradiction. Hence we can infer the truth of the theorem.

We note here that our Legendre condition for the separated problem is

$$(2.21a)\quad [m_{\dot{r}^i \dot{r}^j}(t,y,\dot{y}) + n_{\dot{q}^i \dot{q}^j}(t+\tau,y,\dot{y})]\pi^i\pi^j \geq 0, \quad a \leq t \leq b - \tau;$$

$$(2.21b)\quad m_{\dot{r}^i \dot{r}^j}(t,y,\dot{y})\pi^i\pi^j \geq 0, \quad b - \tau \leq t \leq b,$$

with $\dot{y}$ interpreted in the sense of Convention 2.1.1.

In his book Qualitative Methods of Mathematical Analysis, [3], L. E. El'sgol'c studies the following problem: to minimize

$$J = \int_a^b f[t,y(t-\tau_1(t)),y(t-\tau_2(t)),\ldots,y(t-\tau_n(t)),\dot{y}(t-\tau_1(t)),\ldots,\dot{y}(t-\tau_n(t))]dt$$

on the class of all PWS functions y such that $y(a) = \alpha$, $y(a-m) = \beta$, $y(b-m) = \sigma$, $y(b) = \gamma$ ($\alpha,\beta,\gamma,\sigma$ all constants) where $m = \max_i |\tau_i(a)|$. Also $\tau_i'(t)$ is assumed to be non-negative and such that $1 - \tau_i'(t)$ is bounded away from zero.

If f is written as $f(t,y_1,\ldots,y_n,\dot{y}_1,\ldots,\dot{y}_n)$, then El'sgol'c states without exhibiting any details that the Legendre necessary condition for his problem is

$$\det |f_{\dot{y}_i\dot{y}_k}| \geq 0 \quad (i,k = 1,\ldots,n) \quad \text{for} \quad a \leq t \leq b - m.$$

That this is not the case for our problem will be shown by the following example.

Let y be one-dimensional and consider the separated problem with

$$m(t,y,\mathring{y}) = \mathring{y}^2; \quad n(t,x,\mathring{x}) = -\frac{1}{2}\mathring{x}^2.$$

Then

$$J(y) = \int_a^b [\mathring{y}^2(t) - \frac{1}{2}\mathring{x}^2(t)]dt.$$

Suppose further that $y(t) \equiv 0$, $a - \tau \leq t \leq a$; $y(b) = 0$. It is seen that $J(y)$ is positive definite by writing it in the form

$$J(y) = \int_a^{b-\tau} \frac{1}{2}\mathring{y}^2(t)dt + \int_{b-\tau}^b \mathring{y}^2(t)dt.$$

Then $y(t) \equiv 0$ is a minimizing function and this function clearly satisfies the Euler equations (2.14a) and (2.14b):

$$2\mathring{y}(t) - \mathring{y}(t) = c, \quad a \leq t \leq b - \tau,$$

$$2\mathring{y}(t) = c, \qquad b - \tau \leq t \leq b.$$

The E-function defined by (2.22a) and (2.22b) in the section to follow is identically zero. Also relations(2.21a) and (2.21b) are satisfied since $2 - 1 = 1 > 0$ and $2 > 0$ where

$$g_{rr}(t,y,\mathring{y}) = 2; \quad h_{qq}(t+\tau,y,\mathring{y}) = -1.$$

Now the El'sgol'c result applied to our problem is certainly not true since $g_{rr}h_{qq} = 2(-1) = -2 < 0$.

It is to be noted that the El'sgol'c formulation is distinct from the formulation of the present example so that this example does not bear on the validity of El'sgol'c theorem in [3].

2.4 Weierstrass Necessary Condition

Define a function $E : [a,b] \times R^7 \to R$ by the equations

$$E(t,x,y,z,\mathring{x},\mathring{y},\mathring{z},p) = f(t,x,y,\mathring{x},p) + f(t+\tau,y,z,p,\mathring{z}) - f(t,x,y,\mathring{x},\mathring{y}) - f(t+\tau,y,z,\mathring{y},\mathring{z}) - (p^i-\mathring{y}^i)[f_{r^i}(t,x,y,\mathring{x},\mathring{y}) + f_{q^i}(t+\tau,y,z,\mathring{y},\mathring{z})], \qquad a \le t \le b-\tau; \tag{2.22a}$$

$$E(t,x,y,z,\mathring{x},\mathring{y},\mathring{z},p) = f(t,x,y,\mathring{x},p) - f(t,x,y,\mathring{x},\mathring{y}) - (p^i-\mathring{y}^i)f_{r^i}(t,x,y,\mathring{x},\mathring{y}) \qquad b-\tau \le t \le b. \tag{2.22b}$$

THEOREM 2.4.1 *If* y *furnishes* J *with a strong local minimum then at each* t *in* $[a,b]$,

$$E(t,x(t),y(t),z(t),\mathring{x}(t),\mathring{y}(t),\mathring{z}(t),p) \geq 0 \tag{2.23}$$

in the sense of Convention 2.1.1 *for all real numbers* p.

Let α be a point of the open interval $(a,b-\tau)$ such that α, $\alpha-\tau$, $\alpha+\tau$ do not correspond to corners of y. Choose β such that $\beta < b-\tau$, $0 < \beta-\alpha < \tau$, and such that no corner of y appears for t on the intervals $\alpha-\tau < t < \beta-\tau$, $\alpha < t < \beta$, $\alpha+\tau < t < \beta+\tau$. Now let $Y(t) = (Y^i(t))$ be a function such that $\mathring{Y}(\alpha) \neq \mathring{y}(\alpha)$ but otherwise arbitrary. Here and elsewhere in this proof we understand the symbol $\mathring{Y}(\alpha)$ to mean the right derivative of $Y(t)$ at $t = \alpha$.

Let σ be a real number such that $\alpha < \sigma < \beta$ and define a function $\phi(t,\sigma) = (\phi^i(t,\sigma))$ $(i = 1,\ldots,n)$ by the formulas

$$\phi^i(t,\sigma) = y^i(t) + \frac{Y^i(\sigma) - y^i(\sigma)}{\beta - \sigma} (\beta - t), \quad (i = 1,\dots,n).$$

Let $w(t) = \left(w^i(t)\right)$ be the function defined by the formulas

$$w^i(t) = \begin{cases} y^i(t), & a - \tau \le t < \alpha, \\ Y^i(t), & \alpha \le t < \sigma, \\ \phi^i(t,\sigma), & \sigma \le t < \beta, \\ y^i(t), & \beta \le t \le b, \quad i = 1,\dots,n. \end{cases}$$

Since y minimizes J, we have that $J(w) - J(y) \ge 0$. Set $G(\sigma) = J(w) - J(y)$. Then

$$G(\sigma) = \int_\alpha^\sigma f(t,x,Y,\dot{x},\dot{Y})dt + \int_\sigma^\beta f\left(t,x,\phi(t,\sigma),\dot{x},\phi_t(t,\sigma)\right)dt + \int_{\alpha+\tau}^{\sigma+\tau} f(t,X,y,\dot{X},\dot{y})dt$$

$$+ \int_{\sigma+\tau}^{\beta+\tau} f(t,\phi(t-\tau,\sigma),y,\phi_t(t-\tau,\sigma),\dot{y})dt - \int_\alpha^\beta f(t,x,y,\dot{x},\dot{y})dt - \int_{\alpha+\tau}^{\beta+\tau} f(t,x,y,\dot{x},\dot{y})dt,$$

where $X(t) = Y(t - \tau)$. Upon the linear change of variable $t = s + \tau$ in the third, fourth and sixth integrals, we find that

$$G(\sigma) = \int_\alpha^\sigma [f,t,x,Y,\dot{x},\dot{Y}) + f(t+\tau,Y,z,\dot{Y},\dot{z})dt + \int_\sigma^\beta [f\left(t,x,\phi(t,\sigma),\dot{x},\phi_t(t,\sigma)\right)$$

$$+ f(t+\tau,\phi(t,\sigma),z,\phi_t(t,\sigma),\dot{z})]dt - \int_\alpha^\beta [f(t,x,y,\dot{x},\dot{y}) + f(t+\tau,y,z,\dot{y},\dot{z})]dt$$

Now

$$G'(\sigma) = f\left(\sigma,x(\sigma),Y(\sigma),\dot{x}(\sigma),\dot{Y}(\sigma)\right) + f\left(\sigma+\tau,Y(\sigma),z(\sigma),\dot{Y}(\sigma),\dot{z}(\sigma)\right)$$

$$- f\left(\sigma,x(\sigma),\phi(\sigma,\sigma),\dot{x}(\sigma),\phi_t(\sigma,\sigma)\right) - f\left(\sigma+\tau,\phi(\sigma,\sigma),z(\sigma),\phi_t(\sigma,\sigma),\dot{z}(\sigma)\right)$$

$$+ \int_\sigma^\beta [f_{y^i}(t,x,\phi,\mathring{x},\phi_t)\phi^i_\sigma(t,\sigma) + f_{x^i}(t+\tau,\phi,z,\phi_t,\mathring{z})\phi^i_\sigma(t,\sigma)$$

$$+ f_{r^i}(t,x,\phi,\mathring{x},\phi_t)\phi^i_{t\sigma}(t,\sigma) + f_{q^i}(t+\tau,\phi,z,\phi_t,\mathring{z})\phi^i_{t\sigma}(t,\sigma)]dt.$$

After integrating by parts and setting $\sigma = \alpha$, we find that

$$G'(\alpha) = f\big(\alpha,x(\alpha),y(\alpha),\mathring{x}(\alpha),\mathring{Y}(\alpha)\big) + f\big(\alpha+\tau,y(\alpha),z(\alpha),\mathring{Y}(\alpha),\mathring{z}(\alpha)\big)$$

$$- f\big(\alpha,x(\alpha),y(\alpha),\mathring{x}(\alpha),\mathring{y}(\alpha)\big) - f\big(\alpha+\tau,y(\alpha),z(\alpha),\mathring{y}(\alpha),\mathring{z}(\alpha)\big)$$

$$+ \phi^i_\sigma(t,\alpha)[f_{r^i}\big(t,x(t),y(t),\mathring{x}(t),\mathring{y}(t)\big) + f_{q^i}\big(t+\tau,y(t),z(t),\mathring{y}(t),\mathring{z}(t)\big)]\Big|_\alpha^\beta$$

$$+ \int_\alpha^\beta \Big\{f_{y^i}(t,x,y,\mathring{x},\mathring{y}) + f_{x^i}(t+\tau,y,z,\mathring{y},\mathring{z}) - \frac{d}{dt}[f_{r^i}(t,x,y,\mathring{x},\mathring{y})$$

$$+ f_{q^i}(t+\tau,y,z,\mathring{y},\mathring{z})]\Big\}\phi^i_\sigma(t,\alpha)dt \geq 0.$$

The expression under the integral is zero since y is a minimizing function and hence satisfies the Euler equation. Also $\phi_\sigma(\beta,\alpha) = 0$ so we have the conclusion that (2.23) holds for all t in the open interval $(a,b-\tau)$ where $\mathring{x},\mathring{y},\mathring{z}$ all exist. If t corresponds to a corner of y or if $t = a$ or $t = b - \tau$, use the continuity of f to find that (2.23) holds for all t in $[a,b-\tau]$ in the sense of Convention 2.1.1.

By a similar argument (2.23) holds on $[b-\tau,b]$ with E defined by (2.22b). Note that at $t = b - \tau$ there are two conditions.

We will say that if y_0 satisfies condition (2.23), then it satisfies condition II. The function y_0 satisfies condition II_N if there is a neighborhood $\mathcal{N}$ of the elements $(t,x_0,y_0,\mathring{x}_0,\mathring{y}_0)$ associated with the function y_0 such that the inequality

$$E(t,x,y,z,\mathring{x},\mathring{y},\mathring{z},p) \geq 0$$

holds for all $(t,x,y,\mathring{x},\mathring{y})$ in $\mathscr{N}$ and $(t,x,y,\mathring{x},p)$ in $\mathscr{R}$.

The following result is one which will be useful in proving sufficient conditions.

THEOREM 2.4.2 Suppose that y_0 is an admissible function which satisfies condition II_N. Also suppose that the matrices $|f_{r^ir^j}(t,x,y,\mathring{x},\mathring{y}) + f_{q^iq^j}(t+\tau,y,z,\mathring{y},\mathring{z})|$, $a \leq t \leq b - \tau$ and $|f_{r^ir^j}(t,x,y,\mathring{x},\mathring{y})|$, $b - \tau \leq t \leq b$ are non-singular along y_0. Then there exists a constant $h > 0$ such that

$$E(t,x,y,z,\mathring{x},\mathring{y},\mathring{z},p) \geq h\lambda(p - \mathring{y}) \tag{2.24}$$

for $(t,x,y,\mathring{x},\mathring{y})$ in a neighborhood $\mathscr{N}$ of $(t,x_0,y_0,\mathring{x}_0,\mathring{y}_0)$ and $(t,x,y,\mathring{x},p)$ in $\mathscr{R}$ where

$$\lambda(p - \mathring{y}) = (1 + |p - \mathring{y}|^2)^{\frac{1}{2}} - 1.$$

This theorem has been proved for a Bolza problem with no time-lags by Reid [14]. Also Hestenes [6, p.151] has given a proof for a simpler problem than that considered by Reid. We show that Hestenes' proof may be extended to cover the present problem with delayed argument.

Denote the E-function by the abbreviated symbol $E(\mathring{y},p)$. Now on $[a,b - \tau]$

$$E(\mathring{y},p) = \frac{(\mathring{y}^i-p^i)(\mathring{y}^j-p^j)}{2}[f_{r^ir^j}(t,x,y,\mathring{x},\mathring{y}+\theta(p-\mathring{y})) + f_{q^iq^j}(t+\tau,y,z,\mathring{y}+\theta(p-\mathring{y}),z)].$$

On $[b - \tau,b]$, $E(\mathring{y},p) = \frac{(\mathring{y}^i-p^i)(\mathring{y}^j-p^j)}{2}f_{r^ir^j}(t,x,y,\mathring{x},\mathring{y}+\theta(p-\mathring{y}))$. Then there exists $\mu > 0$ such that $E(\mathring{y},p) \geq \mu|\mathring{y}-p|^2$ on a neighborhood R_1 of y_0. Since

$$|p-\mathring{y}|^2 = [\lambda(p-\mathring{y}) + 2][\lambda(p-\mathring{y})] \geq 2\lambda(p-\mathring{y}),$$

we have that

(2.25) $$E(\mathring{y},p) \geq 2\mu\lambda(p-\mathring{y}).$$

Reduce R_1 so that condition II_N is satisfied with $(t,x,y,\mathring{x},\mathring{y})$ in R_1. Let N_0 be another neighborhood of y_0 such that the closure of N_0 is in R_1 and let $\varepsilon > 0$ be a constant such that if $(t,x,y,\mathring{x},\mathring{y})$ is in N_0, then $(t,x,y,\mathring{x},\mathring{y}+\pi)$ is in R_1 for every vector π with $|\pi| < \varepsilon$. Let $(t,x,y,\mathring{x},\mathring{y})$ be in N_0 and $(t,x,y,\mathring{x},p)$ be in $\mathcal{R} - R_1$. Choose a vector π with $|\pi| = \varepsilon$ and a constant k such that $p = \mathring{y} + k\pi$. Now $k > 1$. From the identity

$$E(\mathring{y},\mathring{y}+\pi) = E(\mathring{y}+\pi,\mathring{y}+k\pi) + E(\mathring{y},\mathring{y}+\pi)k + (k-1)E(\mathring{y}+\pi,\mathring{y})$$

we find that

(2.26) $$E(\mathring{y},\mathring{y}+k\pi) \geq kE(\mathring{y},\mathring{y}+\pi).$$

Now (2.25) holds so that (2.26) becomes

(2.27) $$E(\mathring{y},\mathring{y}+k\pi) \geq 2k\mu\lambda(\pi), \quad k > 1.$$

For $|\pi| = \varepsilon$,

$$\frac{\lambda(\pi)}{\lambda(k\pi)} \geq \frac{\sqrt{1+\varepsilon^2} - 1}{\sqrt{1+k^2\varepsilon^2} - 1} \geq \frac{\varepsilon}{k(2+\pi)}.$$

Now

$$\lambda(\pi) \geq \lambda(k\pi)\frac{\varepsilon}{k(2+\varepsilon)},$$

so by inequality (2.27)

$$E(\mathring{y},\mathring{y}+k\pi) \geq \lambda(k\pi)\frac{2\varepsilon\tau}{(2+\varepsilon)}.$$

Let $h = \frac{2\varepsilon\tau}{2+\varepsilon}$ and $p = \mathring{y} + k\pi$ to get inequality (2.24).

2.5 Examples

EXAMPLE 2.5.1. Let

$$J(y) = \int_0^3 \left(\mathring{y}^2(t) - \mathring{x}^2(t)\right)^2 dt$$

and seek a minimizing function on the class of all PWS functions $y : y(t) = -t, \ t \text{ in } [-1,0], \ y(3) = 1, \ \tau = 1.$ Immediately it is seen that a function with alternate slopes of plus and minus one is a minimizing function. We note that corners can occur anywhere on $[0,3]$.

EXAMPLE 2.5.2. We now produce a similar example which is not quite so trivial in order to apply a future sufficiency criterion. Let

$$J(y) = \int_0^3 \left(\mathring{y}(t) - \dot{x}(t)\right)^2 dt$$

and seek a minimizing function on the class of all PWS functions $y : y(t) = -t, \ -1 < t < 0, \ y(3) = 2 \ \ \tau = 1.$ The Euler equations which a minimizing function must satisfy are the following:

$$2\mathring{y}(t) + \mathring{y}(t+1) + \mathring{y}(t-1) = c, \quad 0 \le t \le 2,$$

$$\mathring{y}(t) + \mathring{y}(t-1) = c, \quad 2 \le t \le 3.$$

A solution of these equations with $c = \frac{1}{2}$ is

$$y(t) = \begin{cases} -t & -1 \le t \le 0, \\ \frac{3}{2}t & 0 < t \le 1, \\ -\frac{3}{2}t+3, & 1 < t \le 2, \\ 2t-4 & 2 < t \le 3. \end{cases}$$

That this is a solution can easily be seen by the following substitution:

$$\text{On } [0,1], \quad 2\left(\frac{3}{2}\right) - \frac{3}{2} - 1 = \frac{1}{2}$$

$$\text{On } [1,2], \quad 2\left(-\frac{3}{2}\right) + 2 + \frac{3}{2} = \frac{1}{2}$$

$$\text{On } [2,3], \quad 2 - \frac{3}{2} = \frac{1}{2}$$

It will be shown later that such a function y actually minimizes J.

It is to be noted that $f_{rr} + f_{qq} = 4 > 0$ and that the minimizing function has a corner at $t = 1$. This shows that Theorem 2.3.2 cannot be extended to cover the case of a general function $f(t,x,y,\dot{x},\dot{y})$.

2.6 A Fourth Necessary Condition

Recall from (2.15) that under the hypothesis that y minimizes the functional J, we have the relation

$$F''(0) = \int_a^b [f_{x^ix^j}\xi^i\xi^j + 2f_{x^iy^j}\xi^i\eta^j + 2f_{x^iq^j}\xi^i\dot{\xi}^j + 2f_{x^ir^j}\xi^i\dot{\eta}^j$$
$$+ f_{y^iy^j}\eta^i\eta^j + 2f_{y^iq^j}\eta^i\dot{\xi}^j + 2f_{y^ir^j}\eta^i\dot{\eta}^j + f_{q^iq^j}\dot{\xi}^i\dot{\xi}^j$$
$$+ 2f_{q^ir^j}\dot{\xi}^i\dot{\eta}^j + f_{r^ir^j}\dot{\eta}^i\dot{\eta}^j]dt \geq 0.$$

The arguments of the partial derivatives of f are $(t,x(t),y(t),\dot{x}(t),\dot{y}(t))$. Denote the integrand by $2\omega(t,\xi,\eta,\dot{\xi},\dot{\eta})$ and consider the integral

$$J(\eta,\lambda) = \int_a^b [2\omega(t,\xi,\eta,\dot{\xi},\dot{\eta}) - \lambda\eta^i\eta^i]dt.$$

In view of Euler's theorem on homogeneous functions this may be written as

$$(2.28) \qquad J(\eta,\lambda) = \int_a^b [\xi^i\omega_{\xi^i} + \eta^i\omega_{\eta^i} + \dot{\xi}^i\omega_{\dot{\xi}^i} + \dot{\eta}^i\omega_{\dot{\eta}^i} - \lambda\eta^i\eta^i]dt.$$

A linear change of variable in the first and third terms of the integrand in (2.24) yields

$$J(\eta,\lambda) = \int_a^b [\theta_i(t)\eta^i(t) + \chi_i(t)\dot{\eta}^i(t) - \lambda\eta^i(t)\eta^i(t)]dt \tag{2.29}$$

where the functions $\theta_i(t)$ and $\chi_i(t)$ are defined as follows:

$$\theta_i(t) = \begin{cases} \omega_{\xi^i}(t+\tau,\eta,\zeta,\dot{\eta},\dot{\zeta}) + \omega_{\eta^i}(t,\xi,\eta,\dot{\xi},\dot{\eta}), & a \leq t \leq b-\tau, \\ \omega_{\eta^i}(t,\xi,\eta,\dot{\xi},\dot{\eta}), & b-\tau < t \leq b; \end{cases}$$

$$\chi_i(t) = \begin{cases} \omega_{\dot{\xi}^i}(t+\tau,\eta,\zeta,\dot{\eta},\dot{\zeta}) + \omega_{\dot{\eta}^i}(t,\xi,\eta,\dot{\xi},\dot{\eta}), & a \leq t \leq b-\tau, \\ \omega_{\dot{\eta}^i}(t,\xi,\eta,\dot{\xi},\dot{\eta}), & b-\tau < t \leq b. \end{cases}$$

The function $\zeta(t)$ is defined by the equation $\zeta(t) \equiv \eta(t+\tau)$.

Note that if $J(\eta,\lambda)$ has a minimum for fixed λ, then by Corollary 2.2.1 it is necessary that the minimizing η satisfy the following equations at points t not corresponding to corners:

$$\omega_{\eta^i}(t) + \omega_{\xi^i}(t+\tau) - \frac{d}{dt}[\omega_{\dot{\eta}^i}(t) + \omega_{\dot{\xi}^i}(t+\tau)] - \lambda\eta^i = 0, \quad a \leq t \leq b-\tau,\ \eta(t) \equiv 0,\ a-\tau \leq t \leq a; \tag{2.30a}$$

$$\omega_{\eta^i}(t) - \frac{d}{dt}\omega_{\dot{\eta}^i}(t) - \lambda\eta^i = 0, \qquad b-\tau \leq t \leq b,\ \eta(b) = 0. \tag{2.30b}$$

DEFINITION 2.6.1 A value λ_0 is said to be a <u>proper value</u> of (2.30a), (2.30b) if there is a non-identically zero solution η_0 of (2.30a) and (2.30b) with $\lambda = \lambda_0$ satisfying the given boundary conditions. Such a solution is called a <u>proper function</u> corresponding to the proper value λ_0.

We assume that proper values of (2.30a), (2.30b) do exist and that they are real.

LEMMA 2.6.1 *If* η_0 *is a proper function corresponding to a proper value* λ_0 *in* (2.30a), (2.30b) *then* $J(\eta_0,\lambda_0) = 0$.

Recall that $J(\eta_0,\lambda_0)$ may be written in the form (2.29). Integrating the terms involving $\dot{\eta}_0^i$ by parts and using the conditions $\eta_0^i(a) = 0 = \eta_0^i(b)$, we obtain the equation

$$J(\eta_0,\lambda_0) = \int_a^b [\theta_i(t) - \frac{d}{dt}\chi_i(t) - \lambda_0\eta_0^i(t)]\eta_0^i(t)dt.$$

The expression in square brackets is just the left-hand side of (2.30a), (2.30b) which η_0 satisfies with λ_0. Hence $J(\eta_0,\lambda_0) = 0$.

We can now prove the following necessary condition.

THEOREM 2.6.1 *If* y *minimizes the functional* J, *there exists no proper value* $\lambda < 0$ *of* (2.30a) *and* (2.30b).

Suppose $\lambda < 0$ is a proper value with proper function η. By the preceding lemma we have the conclusion that

$$F''(0) = \int_a^b \lambda\eta^i\eta^i dt. \tag{2.31}$$

Now $F''(0) \geq 0$ which implies that $\lambda \geq 0$. But this is a contradiction, and hence the theorem is proved.

A complete discussion of a boundary value problem arising from a Bolza problem with no delays is given by Reid [12]. In [13] one finds for the problem of Mayer the relation of a fourth necessary condition involving proper values to a fourth necessary condition using the idea of conjugate points along a solution of Jacobi's equation. Although the present author has not been able to explicitly show that a conjugate point formulation is not applicable to problems with delayed argument, he suspects that such is the case.

CHAPTER III

SUFFICIENT CONDITIONS FOR THE SIMPLE INTEGRAL PROBLEM

3.1 Introduction

In this chapter we give two conditions sufficient to ensure that y_0 in Ω minimizes $J(y)$ on Ω. The first is for global minima patterned after a similar theorem of Ewing [4]; the second for strong local minima is proved by indirect methods after Hestenes [6, p.161] and his student, E. H. Mookini [8, 9].

3.2 A Sufficient Condition for a Global Minimum

There are few known criteria even for classical no-lag problems that establish for a function y_0 that it furnishes a global extremum. Yet one would ultimately desire to have this information. The results of this section like those in [4] are effective for a limited class of problems and yet one which includes many examples with convex integrands that are to be found in the recent literature on control theory.

Define a function $G(t,x_0,x,y_0,y,\mathring{x}_0,\mathring{x},\mathring{y}_0,\mathring{y})$ by the formula

$$G = f - f_0 - (x^i - x_0^i)f_{0x^i} - (y^i - y_0^i)f_{0y^i} - (\mathring{x}^i - \mathring{x}_0^i)f_{0q^i} - (\mathring{y}^i - \mathring{y}_0^i)f_{0r^i}, \quad a \le t \le b, \tag{3.1}$$

where

$$f = f(t,x,y,\mathring{x},\mathring{y}); \quad f_0 = f(t,x_0,y_0,\mathring{x}_0,\mathring{y}_0). \tag{3.2}$$

THEOREM 3.2.1 *Suppose that* y_0 *in* Ω *satisfies equations* (2.7a), (2.7b), *and* (2.8). *Let* y *be an arbitrary function in* Ω. *Then*

$$J(y) - J(y_0) = \int_a^b G\,dt. \tag{3.3}$$

In order to prove this result define the integral

$$\hat{J}(y) = \int_a^b [f_0+(x^i-x_0^i)f_{0x^i}+(\mathring{x}^i-\mathring{x}_0^i)f_{0q^i}+(y^i-y_0^i)f_{0y^i}+(\mathring{y}^i-\mathring{y}_0^i)f_{0r^i}]dt$$

where f_0 is as in (3.2). Using a linear change of variable in the second and third terms, we have that

$$\hat{J}(y) = \int_a^{b-\tau}\{f_0+(y^i-y_0^i)[f_{0y^i}+f_{0x^i}(+\tau)]+(\mathring{y}^i-\mathring{y}_0^i)[f_{0r^i}+f_{0q^i}(+\tau)]\}dt$$
$$+\int_{b-\tau}^b [f_0+(y^i-y_0^i)f_{0y^i}+(\mathring{y}^i-\mathring{y}_0^i)f_{0r^i}]dt$$

where $(+\tau)$ again denotes the set of arguments $(t+\tau, y_0, z_0, \mathring{y}_0, \mathring{z}_0)$. Upon an integration by parts we find that

$$\hat{J}(y) = \int_a^b f_0\,dt+\int_a^{b-\tau}(y^i-y_0^i)[f_{0y^i}+f_{0x^i}(+\tau)-\frac{d}{dt}(f_{0r^i}+f_{0q^i}(+\tau))]dt \tag{3.4}$$
$$+\int_{b-\tau}^b (y^i-y_0^i)(f_{0y^i}-\frac{d}{dt}f_{0r^i})dt+(y^i-y_0^i)[f_{0r^i}+f_{0q^i}(+\tau)]\Big|_a^{b-\tau}$$
$$+(y^i-y_0^i)f_{0r^i}\Big|_{b-\tau}^b.$$

Since both y and y_0 are in Ω, we have that $y(a) = y_0(a)$ and $y(b) = y_0(b)$. Using this fact in conjunction with (2.7a), (2.7b), (2.8) and (3.4), we find that

$$\hat{J}(y) = \int_a^b f(t,x_0,y_0,\mathring{x}_0,\mathring{y}_0)dt.$$

Consequently $\hat{J}(y) = J(y_0)$, so that we have the relation

$$J(y) - J(y_0) = J(y) - \hat{J}(y) = \int_a^b G\, dt.$$

This is conclusion (3.3).

COROLLARY 3.2.1 _If the relation_

$$G(t,x_0,x,y_0,y,\mathring{x}_0,\mathring{x},\mathring{y}_0,\mathring{y}) \geq 0$$

holds for y_0 _satisfying_ (2.7a), (2.7b), _and_ (2.8), _and if_ y _is an arbitrary function in_ Ω, _then_ $J(y_0)$ _is a global minimum for_ $J(y)$.

The proof is immediate from Theorem 3.2.1.

EXAMPLE 3.2.1 Recall Example 2.5.2: Minimize

$$J(y) = \int_0^3 [\mathring{y}(t) \div \mathring{x}(t)]^2 dt$$

in the class of all PWS functions $y : y(t) = -t,\ -1 \leq t \leq 0,\ y(3) = 2,$ $\tau = 1$. Our candidate for the minimizing function is

$$y_0(t) = \begin{cases} -t, & -1 \leq t \leq 0, \\ \frac{3}{2}t, & 0 < t \leq 1, \\ -\frac{3}{2}t+3, & 1 < t \leq 2, \\ 2t-4, & 2 < t \leq 3. \end{cases}$$

Let $p = \mathring{y}$, $q = \mathring{x}$ so that the G-function with our current y_0 becomes the following:

$$\text{On } [0,1], \quad G = (p+q)^2 - (\tfrac{3}{2} - 1)^2 - (p - \tfrac{3}{2})(3-2) - (q+1)(3-2)$$

$$= (p+q)^2 - p - q + \tfrac{1}{4} = [(p+q) - \tfrac{1}{2}]^2 \geq 0;$$

$$\text{On } [1,2], \quad G = (p+q)^2 - (-\tfrac{3}{2} + \tfrac{3}{2})^2 - (p + \tfrac{3}{2})(-3+3) + (q - \tfrac{3}{2})(-3+3)$$

$$= (p+q)^2 \geq 0;$$

$$\text{On } [2,3], \quad G = (p+q)^2 - (2 - \tfrac{3}{2})^2 - (p-2)(4-3) - (q + \tfrac{3}{2})(4-3)$$

$$= (p+q)^2 - p - q + \tfrac{1}{4} = [(p+q) - \tfrac{1}{2}]^2 \geq 0.$$

Hence $G \geq 0$ on $0 \leq t \leq 3$ and for all real numbers p and q. Therefore the corollary to Theorem 3.2.1 yields the result that y_0 furnishes a global minimum.

3.3 Sufficient Conditions for Local Minima by Indirect Methods

This section adapts to certain problems with lags results of M. R. Hestenes and of his student E. H. Mookini for problems without delayed arguments.

In this section we assume that f is free of $\mathring{x}$; i.e. $f = f(t,x,y,r)$. Also we enlarge Ω to be the class of all absolutely continuous vector functions on $[a - \tau, b]$ satisfying $y^i(t) = \alpha^i(t)$, $a - \tau \leq t \leq a$, $y^i(b) = \beta^i$, $i = 1,\ldots,n$ where $\alpha(t) = (\alpha^i(t))$ is a fixed absolutely continuous function with $\mathring{\alpha}(t)$ in $L_2[a - \tau, a]$. Assume that (t,x,y,r) is in the region $\mathcal{R} : [a,b] \times B$ for almost all t in $[a,b]$ where B is an open arcwise connected subset of $R^n \times R^n \times R^n$. Assume also that f and its partial derivatives with respect to x,y, and r evaluated at $(t,x(t),y(t),\mathring{y}(t))$ are integrable on $[a,b]$.

Let y_0 in Ω be a function which satisfies the following conditions:

I (Euler) The function y_0 is of class $C'[a,b]$ (smooth on $[a,b]$) and satisfies the equations

$$f_{r^i}(t) = \int_{t-\tau}^{t} [f_{x^i}(s+\tau) + f_{y^i}(s)ds + c^i, \quad a \le t \le b - \tau;$$

$$f_{r^i}(t) = \int_{b-\tau}^{t} f_{y^i}(s)ds + c^i, \quad b - \tau \le t \le b,$$

where (t) means $\left(t,x_0(t),y_0(t),\mathring{y}_0(t)\right)$ and $(t+\tau)$ means $\left(t+\tau,y_0(t),z_0(t),\mathring{z}_0(t)\right)$.

II_N (Weierstrass) There is a neighborhood N of the elements $(t,x_0,y_0,\mathring{y}_0)$ such that the inequality

$$E(t,x,y,r,p) = f(t,x,y,p) - f(t,x,y,r) - (p^i - r^i)f_{r^i}(t,x,y,r) \ge 0$$

holds for all (t,x,y,r) in N and (t,x,y,p) in $\mathcal{R}$.

III' (Legendre-Clebsch) Along y_0 the following inequality holds for all constant vectors $\pi = (\pi^i) \neq 0$, $i = 1,\ldots,n$;

$$f_{r^i r^j}(t,x_0,y_0,\mathring{y}_0)\pi^i\pi^j > 0.$$

IV' The second variation $J_2(\eta)$ (called $F''(0)$ in Chapter 2) is positive for all $\eta \not\equiv 0$ such that $\eta(t)$ is absolutely continuous and $\mathring{\eta}(t)$ is in $L_2[a-\tau,b]$ and such that $\eta(t) \equiv 0$, $a - \tau < t < a$, $\eta(b) = 0$.

THEOREM 3.3.1 *If y_0 in Ω satisfies the conditions I, II_N, III', and IV' above, then y_0 minimizes the functional*

$$J(y) = \int_a^b f(t,x,y,\mathring{y})dt$$

<u>in the sense that for all</u> y <u>in</u> Ω <u>in a strong neighborhood of</u> y_0,

$$J(y) > J(y_0).$$

The proof of Theorem 3.3.1 will be given after a few preliminary theorems have been noted.

Define the integral $I^*(y)$ by the relation

$$I^*(y) = J(y) - E^*(\mathring{y},\mathring{y}_0)$$

where

$$E^*(\mathring{y},\mathring{y}_0) = \int_a^b E(t,x,y,\mathring{y}_0,\mathring{y})dt.$$

The function $E(t,x,y,\mathring{y}_0,\mathring{y})$ is defined by the relation

(3.5) $$E(t,x,y,\mathring{y}_0,\mathring{y}) = f(t,x,y,\mathring{y}) - f(t,x,y,\mathring{y}_0) - (\mathring{y}^i - \mathring{y}_0^i)f_{r^i}(t,x,y,\mathring{y}_0).$$

Relation (3.5) is of course relations (2.22) appropriately modified in view of the fact that f is free of $\mathring{x}$.

The proofs of the following two theorems are given in [8, p.22,23] for a problem with no delays. Since the proofs are exactly the same, we omit them here.

THEOREM 3.3.2 <u>If</u> y_0 <u>in</u> Ω <u>satisfies condition</u> II_N, <u>then for every</u> $\varepsilon > 0$ <u>there exists a strong neighborhood of</u> y_0 <u>such that</u>

$$|I^*(y) - I^*(y_0)| < \varepsilon[1 + E^*(\mathring{y},\mathring{y}_0)]$$

for every admissible function y in that neighborhood.

THEOREM 3.3.3 *Given a constant* $\sigma > 0$, *there exists* $\rho > 0$ *and a strong neighborhood* F_0 *of* y_0 *such that for every* y *in* F_0

$$J(y) > J(y_0) - \sigma. \tag{3.6}$$

If $E^*(\mathring{y},\mathring{y}_0) \le \rho$, *then*

$$J(y) < J(y_0) + \sigma. \tag{3.7}$$

If $E^*(\mathring{y},\mathring{y}_0) \ge 2\sigma$, *then*

$$J(y) > J(y_0) + \sigma. \tag{3.8}$$

Define the function $K(y,y_0)$ by the formula

$$K(y,y_0) \equiv \int_a^b \lambda(\mathring{y} - \mathring{y}_0)dt \tag{3.9}$$

where

$$\lambda(\mathring{y} - \mathring{y}_0) = \sqrt{1 + |\mathring{y} - \mathring{y}_0|^2} - 1.$$

Since y_0 satisfies conditions II_N and III'; there exists $h > 0$ such that

$$E(t,x,y,z,\mathring{x},r,\mathring{z},p) \ge h\lambda(p-r). \tag{3.10}$$

We now note that if $E^*(\mathring{y},\mathring{y}_0) < \varepsilon h$, then $K(y,y_0) < \varepsilon$.

THEOREM 3.3.4 *Let* $\{y_q\}$ *be a sequence of admissible functions with the property that given a strong neighborhood* F *of* y_0, *there exists an* *integer* q_0 *such that* y_q *is in* F *whenever* $q > q_0$. *If* $\limsup_{q\to\infty} J(y_q) < J(y_0)$, *then*

$$\lim_{q\to\infty} K(y_q,y_0) = 0.$$

To prove this theorem let F be a strong neighborhood of y_0 and assume $E^*(\dot{y}_q,\dot{y}_0) \geq 2\sigma$, $\sigma > 0$, $q > q_0$. Then $J(y_q) > J(y_0) + \sigma$ which contradicts the hypothesis. Hence $E^*(\dot{y}_q,\dot{y}_0) \leq 2\sigma < \varepsilon h$ for $\sigma < \frac{\varepsilon h}{2}$. The above remark applies and

$$K(y_q,y_0) < \varepsilon, \quad \text{for all } \varepsilon > 0, \quad q > q_0.$$

This proves the theorem.

Suppose now that y_0 does not minimize $J(y)$. Then for each integer q there exists y_q in Ω such that

$$J(y_q) \leq J(y_0), \quad \rho_s(y_q,y_0) < \frac{1}{q}. \tag{3.11}$$

Hence $\{y_q\}$ converges uniformly to y_0. It may be shown as in [9, Thm.5.3] that $\{\dot{y}_q\}$ converges in measure to $\dot{y}_0$, and hence there exists a subsequence of $\{y_q\}$ (again called $\{y_q\}$) such that $\{\dot{y}_q\}$ converges almost uniformly to $\dot{y}_0$.

THEOREM 3.3.5 *Let* $\{y_q\}$ *be the sequence of functions defined by* (3.11). *Given a constant* $\rho > 0$ *there is a constant* $\delta > 0$ *and an integer* q *such that if* M *is a subset of* [a,b] *of measure at most* δ *and* $q > q_0$, *then*

$$0 \leq \int_M \lambda_q(t)dt < \rho \tag{3.12}$$

<u>where</u>

$$\lambda_q(t) = \lambda(\mathring{y}_q - \mathring{y}_0) + 2. \tag{3.13}$$

The proof of this theorem is also to be found in [9].

Define the constant k_q by the relation

$$k_q^2 = K(y_q, y_0).$$

Let

$$\eta_q(t) = \frac{y_q(t) - y_0(t)}{k_q}.$$

The variation $\eta_q(t)$ satisfies the relation

$$\int_a^b \frac{|\dot{\eta}_q|^2}{\lambda_q(t)}\, dt = 1. \tag{3.14}$$

This relation follows from the identity

$$\lambda(\dot{y}_q - \dot{y}_0) = \frac{|\mathring{y}_q - \mathring{y}_0|^2}{\lambda_q(t)}.$$

The following theorem is proved in [9, Thm. 6.2].

THEOREM 3.3.6 <u>Let</u> $\{y_q\}$ <u>be the sequence of functions defined by</u> (3.11) <u>and let</u> $\mathring{\eta}_q = k_q^{-1}(\mathring{y}_q - \mathring{y}_0)$. <u>There is a function</u> $\dot{\eta}_0(t)$ <u>in</u> $L_2[a-\tau, b]$ <u>such that the sequence</u> $\{\mathring{\eta}_q\}$ <u>has a subsequence (again called</u> $\{\dot{\eta}_q\}$) <u>which converges weakly to</u> $\mathring{\eta}_0(t)$ <u>in</u> $L_2[a-\tau, b]$ <u>on every measurable set</u> M <u>on which</u> $\{\mathring{y}_q(t)\}$ <u>converges uniformly to</u> $\mathring{y}_0(t)$. <u>Moreover, for every bounded integrable function</u> $g(t)$,

$$\lim_{q\to\infty} \int_a^b g(t)\mathring{\eta}_q(t)dt = \int_a^b g(t)\mathring{\eta}_0(t)dt. \tag{3.15}$$

The following two theorems are both proved in [6, p.157].

THEOREM 3.3.7 *Let* $\{y_q\}$ *be the sequence of functions of the last theorem, and let* η_q *be defined as before:*

$$\eta_q(t) = \frac{y_q(t) - y_0(t)}{k_q} .$$

There exists a function $\eta_0(t)$ *whose derivative is* $\dot{\eta}_0(t)$ *such that* $\{\eta_q(t)\}$ *converges uniformly to* $\eta_0(t)$ *on* $[a,b]$.

THEOREM 3.3.8 *Let* $N_{iq}(t)$, $N_i(t)$ *be continuous functions such that*

$$\lim_{q\to\infty} N_{iq}(t) = N_i(t) \quad \text{uniformly on } [a,b];$$

then

$$\lim_{q\to\infty} \int_a^b N_{iq}(t)\dot{\eta}_q^i(t)dt = \int_a^b N_i(t)\dot{\eta}_0^i(t)dt.$$

We now write $J(y)$ in the following way:

$$J(y) = J(y_0) + J_1(y - y_0) + K(y) + E^*(\dot{y},\dot{y}_0)$$

where

$$J_1(y-y_0) = \int_a^b [(x^i-x_0^i)f_{0x^i} + (y^i-y_0^i)f_{0y^i} + (\mathring{y}^i-\mathring{y}_0^i)f_{0r^i}]dt;$$

$$K(y) = \int_a^b [M(t,y) + (\mathring{y}^i-\mathring{y}_0^i)N_i(t,y)]dt.$$

Here

$$M(t,y) = f(t,x,y,\mathring{y}_0) - f(t,x_0,y_0,\mathring{y}_0) - (x^i-x_0^i)f_{0x^i} - (y^i-y_0^i)f_{0y^i};$$

$$N_i(t,y) = f_{r^i}(t,x,y,\mathring{y}_0) - f_{r^i}(t,x_0,y_0,\mathring{y}_0).$$

Now M and N_i can be written as follows:

$$M(t,y) = \int_0^1 (1-\theta)[f_{x^ix^j}(x^i-x_0^i)(x^j-x_0^j)$$

$$+ 2f_{x^iy^j}(x^i-x_0^i)(y^j-y_0^j) + f_{y^iy^j}(y^i-y_0^i)(y^j-y_0^j)]d\theta$$

where the partial derivatives of f are evaluated at $(t,x_0+\theta(x-x_0),y_0+\theta(y-y_0),\mathring{y}_0)$, $0 < \theta < 1$.

$$N_i(t,y) = \int_0^1 [f_{r^ix^j}(x^j-x_0^j) + f_{r^iy^j}(y^j-y_0^j)]d\theta.$$

The partial derivatives of f are evaluated as above.

Recall that

$$I^*(y) = J(y) - E^*(\mathring{y},\mathring{y}_0) \quad \text{and} \quad I^*(y_0) = J(y_0).$$

Using the fact that y_0 satisfies the Euler conditions, we find that

$$(3.16)\quad \lim_{q\to\infty} \frac{I^*(y_q) - I^*(y_0)}{k_q^2} = \lim_{q\to\infty} \frac{1}{k_q^2}\int_a^b [M(t,y_q) + (\mathring{y}^i - \mathring{y}_0^i)N_i(t,y_q)]dt$$

$$= \frac{1}{2}J_2(\eta_0) - \frac{1}{2}\int_a^b f_{r^i r^j}(t,x_0,y_0,\mathring{y}_0)\eta_0^i\eta_0^j dt.$$

We now show that

$$\liminf_{q\to\infty} k_q^{-2}E^*(\mathring{y}_q,\mathring{y}_0) \geq \frac{1}{2}\int_a^b f_{r^i r^j}(t,x_0,y_0,\mathring{y}_0)\dot{\eta}_0^i\dot{\eta}_0^j dt.$$

Let M be a measurable subset of $[a,b]$ on which $\{\mathring{y}_q\}$ converges to $\dot{y}_0$ uniformly. There exists an integer q_0 such that if $q \geq q_0$, then

$$(3.17)\quad 2E(t,x_q,y_q,\mathring{y}_0,\mathring{y}_q) = k_q^2 f_{r^i r^j}\big(t,x_q,y_q,\mathring{y}_0+\theta(\mathring{y}_q-\mathring{y}_0)\big)\dot{\eta}_0^i\dot{\eta}_0^j .$$

Also

$$\lim_{q\to\infty} f_{r^i r^j}\big(t,x_q,y_q,\mathring{y}_0+\theta(\mathring{y}_q-\mathring{y}_0)\big) = f_{r^i r^j}(t,x_0,y_0,\mathring{y}_0),\ \text{uniformly on } M.$$

Let

$$f_{ij}(x,y) = f_{r^i r^j}[t,x,y,\mathring{y}_0+\theta(\mathring{y}-\mathring{y}_0)].$$

Then

$$\int_M f_{ij}(x_q,y_q)\dot{\eta}_q^i\dot{\eta}_q^j dt = \int_M f_{ij}(x_0,y_0)\dot{\eta}_q^i\dot{\eta}_q^j dt + \int_M [f_{ij}(x_q,y_q) - f_{ij}(x_0,y_0)]\dot{\eta}_q^i\dot{\eta}_q^j dt.$$

The last integral has limit zero as $q\to\infty$ so

$$\liminf_{q\to\infty} \int_M f_{ij}(x_q,y_q)\dot{\eta}_q^i\dot{\eta}_q^j = \liminf_{q\to\infty} \int_M f_{ij}(x_0,y_0)\dot{\eta}_q^i\dot{\eta}_q^j .$$

Since $\{\mathring{\eta}_q\}$ converges weakly to $\mathring{\eta}_0$ on M, we have the result that

$$\lim_{q\to\infty} \int_M f_{ij}(x_0,y_0)\mathring{\eta}_0^i\mathring{\eta}_q^j = \int_M f_{ij}(x_0,y_0)\mathring{\eta}_0^i\mathring{\eta}_0^j.$$

Hence

$$\liminf_{q\to\infty} \int_M f_{ij}(x_q,y_q)\mathring{\eta}_q^i\mathring{\eta}_q^j = \int_M f_{ij}(x_0,y_0)\mathring{\eta}_0^i\mathring{\eta}_0^j$$

$$+ \liminf_{q\to\infty} \int_M f_{ij}(x_0,y_0)(\mathring{\eta}_q^i-\mathring{\eta}_0^i)(\mathring{\eta}_q^j-\mathring{\eta}_0^j).$$

Applying III' and (3.17) to the last integral we find that

$$\liminf_{q\to\infty} k_q^{-2}\int_M E(\mathring{y}_q,\mathring{y}\ ) \geq \frac{1}{2}\int_M f_{ij}(x_0,y_0)\mathring{\eta}_0^i\mathring{\eta}_0^j.$$

Using condition II_N for $E(\mathring{y}_q,\mathring{y}_0)$, we see that

$$\liminf_{q\to\infty} k_q^{-2}E^*(\mathring{y}_q,\mathring{y}_0) \geq \frac{1}{2}\int_M f_{ij}(x_0,y_0)\mathring{\eta}_0^i\mathring{\eta}_0^j.$$

Recall that for all $\varepsilon > 0$, M may be chosen so that the measure of M differs from $b - a$ by less than ε. Hence

$$\text{(3.18)} \qquad \liminf_{q\to\infty} k_q^{-2}E^*(\mathring{y}_q,\mathring{y}_0) \geq \frac{1}{2}\int_a^b f_{ij}(x_0,y_0)\mathring{\eta}_0^i\mathring{\eta}_0^j dt.$$

From the definition of $I^*(y)$ we see that

$$\limsup_{q\to\infty} \frac{I^*(y_q) - I^*(y_0)}{k_q^2} \leq \limsup_{q\to\infty} \frac{J(y_q) - J(y_0)}{k_q^2} - \liminf_{q\to\infty} k_q^{-2}E^*(\mathring{y}_q,\mathring{y}_0).$$

Hence we have by relations (3.16), (3.18) that

$$0 \geq \limsup_{q\to\infty} \frac{J(y_q) - J(y_0)}{k_q^2} = \frac{1}{2} J_2(\eta_0) + \liminf_{q\to\infty} k_q^{-2} E^*(\mathring{y}_q, \mathring{y}_0)$$

$$- \frac{1}{2} \int_a^b f_{r^i r^j}(t, x_0, y_0, \mathring{y}_0) \mathring{\eta}_0^i \mathring{\eta}_0^j \geq \frac{1}{2} J_2(\eta_0).$$

It now follows that $\eta_0(t) \equiv 0$ since $J_2(\eta)$ was assumed to be positive definite. We then can conclude that $0 \geq \liminf_{q\to\infty} k_q^{-2} E^*(\mathring{y}_q, \mathring{y}_0)$. This result and (3.10) imply that

$$0 \geq \liminf_{q\to\infty} k_q^{-2} \int_a^b h\lambda(\mathring{y}_q - \mathring{y}_0)dt = h.$$

But $h > 0$, so that the assumption that $J(y_q) \leq J(y_0)$ is false. Theorem 3.3.1 is therefore proved.

CHAPTER IV

THE PROBLEM OF HESTENES WITH DELAYED ARGUMENT

4.1 Preliminary Theorems

In this chapter we adapt certain results on control problems obtained by Hestenes [7; 6, Chap. 6] to control problems with delayed arguments. In particular we extend Theorem 3.1 of [7] to time lag problems. Also we discuss the second variation of such problems. The final section of this chapter contains a sufficient condition for optimality which does not depend on the sign of the second variation but is of rather limited applicability.

Following Hestenes we consider an arc to be a system

$$y(t) = \left(y^1(t),\ldots,y^n(t)\right), \quad u(t) = \left(u^1(t),\ldots,u^q(t)\right).$$

Such an arc will be denoted by the single symbol y. The function $y(t)$ is a continuous (PWS) n-dimensional vector function; the function $u(t)$ is a PWC q-dimensional vector function. We define $x(t) \equiv y(t - \tau)$, $z(t) \equiv y(t + \tau)$, $v(t) = u(t + \tau)$ where τ is a positive constant.

The problem is that of minimizing a functional

$$I_0(y) = \int_a^b L_0\left(t,x(t),y(t),u(t)\right)dt \tag{4.1}$$

in a class of arcs

y: $y(t),\ u(t)$

satisfying a system of differential-difference equations

(4.2) $$\mathring{y}^i(t) = f^i\big(t,x(t),y(t),u(t)\big), \quad i = 1,\dots,n,$$

the set of initial and terminal conditions

(4.3) $$y^i(t) = \alpha^i(t), \quad a - \tau \le t \le a, \quad y^i(b) = \beta^i,$$

and a set of isoperimetric relations

(4.4) $$\begin{aligned} I_\gamma(y) &\le 0, \quad 1 \le \gamma \le p' \\ I_\gamma(y) &= 0, \quad p' < \gamma \le p \end{aligned}$$

where

$$I_\gamma(y) = \int_a^b L_\gamma[t,x(t),y(t),u(t)]dt.$$

The conditions (4.2) and (4.3) in vector notation are

(4.5) $$\mathring{y}(t) = f\big(t,x(t),y(t),u(t)\big)$$

(4.6) $$y(t) = \alpha(t), \quad a - \tau \le t \le a, \quad y(b) = \beta.$$

We assume that the functions f^i, L_0, L_γ are all continuous and have continuous first partial derivatives on a region $\mathcal{R}$ of (t,x,y,u)-space, $[a,b] \times B$ where B is an open arcwise connected subset of $R^n \times R^n \times R^q$. Denote by $\mathcal{R}_0$ a subset of $\mathcal{R}$ with the property that if $(\bar{t},\bar{x},\bar{y},\bar{u})$ is in $\mathcal{R}_0$, then there exists a continuous vector function

$u(t)$ on $\bar{t} - \delta \le t \le \bar{t} + \delta$ such that $u(\bar{t}) = \bar{u}$ and $|t,x,y,u(t)|$ is in $\mathcal{R}_0$ whenever $|t - \bar{t}| \le \delta$, $|x - \bar{x}| \le \delta$, $|y - \bar{y}| \le \delta$. We will denote by $\mathcal{O}$ the class of all arcs y such that $(t,x(t),y(t),u(t))$ is in $\mathcal{R}_0$ along y, such that $y(t)$ is a solution of (4.5), and such that $y(t)$ satisfies (4.6). The vector function $\alpha(t)$ is PWS on $[a - \tau, a]$, and β is a constant vector.

In preparation for Theorem 4.1.1 we make the following definition which is to be found in [7]. Let $J_\rho(y)$, $\rho = 0,1,\ldots,r$ be a set of real valued functions defined on a space $\mathcal{B}$ of elements y. Let y_0 be in $\mathcal{B}$. A set K of vectors $k = (k^0,\ldots,k^n)$ in a Euclidean space R^{n+1} will be called a <u>derived set of vectors for</u> J_ρ <u>at</u> y_0 <u>on</u> $\mathcal{B}$ if given any finite set of vectors $k_1,\ldots,k_N$ in K, there is a function

$$y(h) = y(h_1,\ldots,h_N)$$

defined on a set $0 \le h_j \le \delta$, $j = 1,\ldots,N$, $\delta > 0$, with values in $\mathcal{B}$ such that $y(0) = y_0$ and such that the functions

$$\phi_\rho(h) = J_\rho\big(y(h)\big) - J_\rho(y_0), \quad \rho = 0,1,\ldots,r$$

are continuous on the set $0 \le h_j \le \delta$ and have $d\phi_\rho = k_j^\rho dh_j$ as their differentials at $h = 0$ on the same set.

We now state the following theorem which is proved in [7, p.39].

THEOREM 4.1.1 <u>Suppose that</u> K <u>is a derived set for</u> J_ρ <u>at</u> y_0 <u>on</u> $\mathcal{B}$. <u>If</u> y_0 <u>minimizes</u> $J_0(y)$ <u>on</u> $\mathcal{B}$ <u>subject to the constraints</u>

$$(4.7)\qquad \begin{aligned} &J_\gamma(y) \le 0, \quad 1 \le \gamma \le r' \\ &J_\gamma(y) = 0, \quad r' < \gamma \le r \end{aligned}$$

then there exist multipliers $\lambda_0 \geq 0, \lambda_1,\ldots,\lambda_r$ not all zero, such that the inequality

$$L(k) = \lambda_\rho k^\rho \geq 0, \quad \rho = 0,1,\ldots,r \tag{4.8}$$

holds for every vector k in K and hence for every vector in the closure K^* of the convex cone generated by K. Moreover, $\lambda_\gamma \geq 0$, $1 \leq \gamma \leq r'$ with $\lambda_\gamma = 0$ if $J_\gamma(y_0) < 0$.

4.2 A Maximum Principle

The following theorem extends for problems with delayed argument a result established by Hestenes [7, Thm.3.1]. The results are quite similar to those found in [11, p.213], but the method of proof is quite different. Also isoperimetric relations do not appear explicitly in [11].

THEOREM 4.2.1 Suppose that the arc

$$y_0: \qquad y_0(t),\ u_0(t), \qquad a \leq t \leq b$$

affords a minimum to I_0 on $\mathfrak{A}$. Then there exist multipliers

$$\lambda_0 \geq 0,\ \lambda_\gamma,\ p_i(t), \qquad \gamma = 1,\ldots,p;\ i = 1,\ldots,n$$

not vanishing simultaneously on $[a,b]$ and the function

$$H(t,x,y,u,p) = p_i f^i(t,x,y,u) - \lambda_0 L_0(t,x,y,u) - \lambda_\gamma L_\gamma(t,x,y,u) \tag{4.9}$$

(summed on i and γ) such that the following relations hold.

(i) The multipliers λ_γ are constant and $\lambda_\gamma \geq 0$, $1 \leq \gamma \leq p'$ with $\lambda_\gamma = 0$ in case $I_\gamma(y_0) < 0$.

(ii) The multipliers $p_i(t)$ are continuous and have piecewise continuous derivatives. There are constants d_i such that the relations

$$(4.10)\quad p_i(t) = -\int_{b-\tau}^{t}[H_{y^i}(\ ) + H_{x^i}(+\tau)]ds + d_i, \qquad a \le t \le b-\tau,$$

$$(4.11)\quad p_i(t) = -\int_{b-\tau}^{t} H_{y^i}(\ )ds + d_i, \qquad b-\tau \le t \le b,$$

hold along y_0.

(iii) The inequality

$$(4.12)\quad H(t,x_0(t),y_0(t),u,p(t)) \le H(t,x_0(t),y_0(t),u_0(t),p(t))$$

holds whenever $(t,x_0(t),y_0(t),u)$ is in $\mathcal{R}_0$.

The formula (4.10) is an abbreviation for the expression

$$p_i(t) = -\int_{b-\tau}^{t}\{H_{y^i}[s,x_0(s),y_0(s),u_0(s),p(s)]$$

$$+ H_{x^i}[s+\tau,y_0(s),z_0(s),v_0(s),p(s+\tau)]\}ds + d_i.$$

A similar remark holds for the expression (4.11).

In beginning the proof we make the following definitions. For $i,j = 1,\ldots,n$ and $\gamma = 0,1,\ldots,p$, set

$$A^i_j(t) = f^i_{y^j}(t,x_0(t),y_0(t),u_0(t))$$

$$b_{\gamma j}(t) = L_{\gamma y^j}(t,x_0(t),y_0(t),u_0(t))$$

$$C^i_j(t) = f^i_{x^j}(t,x_0(t),y_0(t),u_0(t))$$

$$d_{\gamma j}(t) = L_{\gamma x^j}(t,x_0(t),y_0(t),u_0(t)).$$

Now let $A(t)$ be the matrix $(A^i_j(t))$. Likewise $C(t) = (C^i_j(t))$. Let

$b_\gamma(t)$ be the n-dimensional vector $(b_{\gamma j}(t))$. Also $d_\gamma(t) = (d_{\gamma j}(t))$.

Let $q_\gamma(t)$ be a solution of the differential-difference system

$$\dot{q}_\gamma(t) = -q_\gamma(t)A(t) - q_\gamma(t+\tau)C(t+\tau) - b_\gamma(t) - d_\gamma(t+\tau), \quad a \le t < b - \tau,$$

$$\dot{q}_\gamma(t) = -q_\gamma(t)A(t) - b_\gamma(t), \qquad b - \tau \le t \le b,$$

with $q_\gamma(b-\tau) = 0$. By applying standard existence theorems for differential equations to the above system on the successive intervals $[b - \tau, b]$, $[b - 2\tau, b - \tau]$,... one finds that a PWS solution to the system does exist.

Now set

$$(4.13) \qquad f_\gamma(t,x,y,u) = L_\gamma(t,x,y,u) + \frac{d}{dt}(q_\gamma y), \quad (\gamma = 0,1,\dots,p)$$

Define the function $J_\gamma(y)$ by the formula

$$J_\gamma(y) = \int_a^b f_\gamma(t,x,y,u)dt + c_\gamma, \quad \gamma = 0,1,\dots,p.$$

where

$$c_\gamma = -q_\gamma(b)y(b) + q_\gamma(a)y(a).$$

It is easily seen that

$$J_\gamma(y) = I_\gamma(y)$$

for y in $\mathfrak{A}$. Hence y_0 minimizes $J_0(y)$ subject to the constraints

$$J_\gamma(y) \le 0, \qquad 1 \le \gamma \le p'$$

$$J_\gamma(y) = 0, \qquad p' < \gamma \le p.$$

Let the nxn-matrix $G(t) = \left(G_{ij}(t)\right)$ be a solution of the matrix differential-difference equations

$$\mathring{G}(t) = -G(t)A(t) - G(t+\tau)C(t+\tau), \quad a \le t < b - \tau,$$

$$\mathring{G}(t) = -G(t)A(t) \qquad\qquad b - \tau \le t \le b,$$

with $G(b) = I$ where I is the nxn identity matrix. As before a PWS solution to this system exists. Define the function $f_{p+i}(t,x,y,u)$ by the formula

$$(4.14) \qquad f_{p+i}(t,x,y,u) = \frac{d}{dt}\left(G_{ij}(t)y^j(t)\right), \quad i = 1,\ldots,n.$$

Set

$$J_{p+i}(y) = \int_a^b f_{p+i}(t,x,y,u)dt - c_i^o$$

where we define

$$c_i^o = \beta^i - G_{ij}(a)y^j(a).$$

It is now clear that y in $\mathcal{O}\!\mathcal{l}$ is equivalent to the condition $J_{p+i}(y) = 0 \ (i = 1,\ldots,n)$ since

$$J_{p+i}(y) = y^i(b) - \beta^i.$$

Let K be all $n+p+1$-vectors of the form $k = (k^\rho)$ where

$$(4.15) \quad k^\rho = f_\rho(t,x_0,y_0,u) - f_\rho(t,x_0,y_0,u_0), \qquad \rho = 0,1,\ldots,p+n$$

where t is in the interval $a < t < b$, and t is neither a point of discontinuity for u_0 nor a point of discontinuity of $\mathring{q}$ or $\dot{G}$. We now have the following.

LEMMA 4.2.1 <u>The class</u> K <u>is a derived set of vectors for</u> J_ρ <u>at</u> y_0 <u>on</u> $\mathcal{O}\!\!l$.

Granting for the moment the truth of the lemma for which the proof will be given later, we proceed with the main line of the proof of Theorem 4.2.1. By Theorem 4.1.1 there exist multipliers $\lambda_0 \geq 0$, $\lambda_1, \ldots, \lambda_{p+n}$, not all zero such that

$$(4.16) \qquad L(k) = \lambda_\rho k^\rho \geq 0, \qquad \rho = 0, 1, \ldots, p+n,$$

for k in K. Also $\lambda_\gamma \geq 0$, $1 \leq \gamma \leq p'$, with $\lambda_\gamma = 0$ if $J_\gamma(y_0) < 0$. Now set

$$(4.17) \qquad F(t,x,y,u) = \lambda_\rho f_\rho(t,x,y,u).$$

We then see using (4.15) and (4.16) that

$$(4.18) \qquad F(t,x_0,y_0,u) \geq F(t,x_0,y_0,u_0)$$

except possibly at a finite number of values of t on $[a,b]$. Theorem 4.1.1 guarantees that (4.16) holds for all k in the closure K^* of the convex cone generated by K. Hence (4.18) holds for all t in $[a,b]$ with (t,x_0,y_0,u) in $\mathcal{R}_0$.

We now define the functions $p_i(t)$ by the equation

$$p_i(t) = -\lambda_\gamma q_{\gamma i}(t) - \lambda_{p+j} G_{ij}(t), \qquad \gamma = 0,1,\ldots,p;\ j = 1,\ldots,n.$$

Using the definition (4.17) of F, the definition of H and (4.18) we find that

$$H(t,x_0,y_0,u,p) \leq H(t,x_0,y_0,u_0,p).$$

Differentiate $p_i(t)$ to find that the following relation holds for all values of t on $[a,b-\tau]$ except those which correspond to discontinuities of u_0, $\mathring{q}_{\gamma i}$, or $\mathring{G}_{ij}$.

$$\mathring{p}_i(t) = -H_{y^i}[t,x_0(t),y_0(t),u_0(t),p(t)]-H_{x^i}[t+\tau,y_0(t),z_0(t),v_0(t),p(t+\tau)].$$

On $[b-\tau,b]$, the relation

$$\mathring{p}_i(t) = -H_{y^i}[t,x_0(t),y_0(t),u_0(t),p(t)]$$

holds except for values of t corresponding to discontinuities of u_0.

PROOF OF LEMMA 4.2.1 Let $k_1,\ldots,k_N$ be N vectors in K. They are of the form $k_j = (k_j^\rho)$ where

$$k_j^\rho = f_\rho(t_j,x_0(t_j),y_0(t_j),u_j)-f_\rho\big(t_j,x_0(t_j),y_0(t_j),u_0(t_j)\big), \qquad i = 1,\ldots,N.$$

We may assume that $t_1 < t_2 < \ldots < t_N$. By assumption there exist functions $u_j(t)$ defined on a ε-neighborhood of t_j such that $\big(t,x_0(t),y_0(t),u_j(t)\big)$ is in $\mathcal{R}_0$ and $u_j(t_j) = u_j$. Choose $\delta > 0$ such that $N\delta < \varepsilon$ and $\delta < \tau$. Also we require that the intervals

$$t_i \leq t \leq t_i + N\delta, \quad t_j \leq t \leq t_j + N\delta$$

be disjoint when $i \neq j$. Let $\mathcal{H}$ be the set of all vectors $h = (h_j)$, $j = 1,\ldots,N$, such that $0 \leq h_j \leq \delta$. Set $T_1 = t_1$, $T_j = t_j + h_1 + \ldots + h_{j-1}$, $j = 2,\ldots,N$. Let $M(h)$ be the complement in $[a,b]$ of the set of intervals $T_j \leq t \leq T_j + h_j$, $j = 1,\ldots,N$. Define a function $u(t,h)$ by the formula

$$u(t,h) = u_j(t), \quad T_j \le t \le T_j + h_j, \quad j = 1,\ldots,N$$

$$u(t,h) = u_0(t), \quad t \text{ in } M(h).$$

If δ' is chosen sufficiently small, it follows from standard theorems, e.g. [5, Chap. IX, Sec. 3], applied to the successive intervals $[a, a+\tau]$, $[a+\tau, a+2\tau], \ldots$ that the differential-difference system

$$\dot{y}^i = f^i\big(t,x,y,u(t,h)\big), \quad y(t) = \alpha(t), \quad a - \tau \le t \le a$$

has a solution $y(t,h)$ for $a \le t \le b$, $0 \le h \le \delta'$. The arc

$$y(h): \qquad y(t,h),\ u(t,h), \quad a \le t \le b$$

has the property that $y(0) = y_0$. Moreover, $y(t,h)$ has partial derivatives with respect to the h_j which are PWC functions of t on $[a,b]$.

The functions

$$\phi_\rho(h) = J_\rho\big(y(h)\big) - J_\rho(y_0)$$

are of class C' on $\mathcal{H}$. Setting $h_i = 0$, $i \ne j$, we see that

$$\phi_\rho(h) = P_\rho(h) + Q_\rho(h)$$

where

$$P_\rho(h) = \int_{T_j}^{T_j+h_j} [f_\rho\big(t,x(t,h),y(t,h),u_j(t)\big) - f_\rho\big(t,x_0(t),y_0(t),u_0(t)\big)]dt\ ;$$

$$Q_\rho(h) = \int_{M(h)} [f_\rho\big(t,x(t,h),y(t,h),u_0(t)\big) - f_\rho\big(t,x_0(t),y_0(t),u_0(t)\big)]dt.$$

Now at $h_j = 0$ we have the result that

$$\frac{\partial P_\rho}{\partial h_j} = k_j^\rho .$$

It remains to show that at $h_j = 0$, $\frac{\partial Q_\rho}{\partial h_j} = 0$. Observe that except for a finite set of points $M(h)$ is the union of the intervals $[a,T_j)$ and $(T_j + h_j, b]$. Hence

$$Q_\rho(h) = \int_a^{T_j} [f_\rho(t,x(t,h),y(t,h),u_0(t)) - f_\rho(t,x_0(t),y_0(t),u_0(t))]dt$$

$$+ \int_{T_j+h_j}^{b} [f_\rho(t,x(t,h),y(t,h),u_0(t)) - f_\rho(t,x_0(t),y_0(t),u_0(t))]dt.$$

Therefore we have that

$$(4.20) \quad \frac{\partial Q_\rho}{\partial h_j} = \int_a^{T_j} \left[\frac{\partial f_\rho}{\partial y^\ell}\frac{\partial y^\ell}{\partial h_j} + \frac{\partial f_\rho}{\partial x^\ell}\frac{\partial x^\ell}{\partial h_j}\right]dt + \int_{T_j+h_j}^{b} \left[\frac{\partial f_\rho}{\partial y^\ell}\frac{\partial y^\ell}{\partial h_j} + \frac{\partial f_\rho}{\partial x^\ell}\frac{\partial x^\ell}{\partial h_j}\right]dt$$

$$- f_\rho[T_j+h_j, x(T_j+h_j,h), y(T_j+h_j,h), u_0(T_j+h_j)]$$

$$+ f_\rho[T_j+h_j, x_0(T_j+h_j,h), y_0(T_j+h_j,h), u_0(T_j+h_j)].$$

Note that at $h_j = 0$, the last two terms add to zero. A linear change of variable yields the result that

$$(4.21) \quad \frac{\partial Q_\rho}{\partial h_j} = \int_a^{T_j-\tau} \left[\frac{\partial f_\rho}{\partial y^\ell}(\;) + \frac{\partial f_\rho}{\partial x^\ell}(+\tau)\right]\frac{\partial y^\ell}{\partial h_j}dt + \int_{T_j-\tau}^{T_j} \frac{\partial f_\rho}{\partial y^\ell}(\;)\frac{\partial y^\ell}{\partial h_j}dt$$

$$+ \int_{T_j+h_j-\tau}^{T_j+h_j} \frac{\partial f_\rho}{\partial x^\ell}(+\tau)\frac{\partial y^\ell}{\partial h_j}dt + \int_{T_j+h_j}^{b-\tau} \left[\frac{\partial f_\rho}{\partial y^\ell}(\;) + \frac{\partial f_\rho}{\partial x^\ell}(+\tau)\right]\frac{\partial y^\ell}{\partial h_j}dt$$

$$+ \int_{b-\tau}^{b} \frac{\partial f_\rho}{\partial y^\ell}(\ ) \frac{\partial y^\ell}{\partial h_j}\, dt.$$

Here $\frac{\partial f_\rho}{\partial y^\ell}(\ )$ denotes the expression $\frac{\partial f_\rho}{\partial y^\ell}\big(t, x(t,h), y(t,h), u_0(t)\big)$; the term $\frac{\partial f_\rho}{\partial x^\ell}(+\tau)$ stands for $\frac{\partial f_\rho}{\partial x^\ell}\big(t+\tau, y(t,h), z(t,h), v_0(t)\big)$. Note that at $h_j = 0$, the second and third integrals of (4.21) become

$$\int_{T_j-\tau}^{T_j} \left[\frac{\partial f_\rho}{\partial y^\ell}(\ ) + \frac{\partial f_\rho}{\partial x^\ell}(+\tau)\right] \frac{\partial y^\ell}{\partial h_j}\, dt.$$

Since $\frac{\partial f_\rho}{\partial y^\ell}(\ ) + \frac{\partial f_\rho}{\partial x^\ell}(+\tau)$ vanishes along y_0 by the definitions (4.13) and (4.14) of f_ρ we have that $\frac{\partial Q_\rho}{\partial h_j}$ vanishes for $j = 1,\ldots,N$. Hence we conclude that

$$\frac{\partial \phi_\rho}{\partial h_j} = k_j^\rho$$

at $h = 0$ for $\rho = 0,1,\ldots,p+n$. We therefore conclude that the differential of $\phi_\rho(h)$ at $h=0$ is $k_j^\rho dh_j$ as was to be proved.

COROLLARY 4.2.1 <u>If</u> $\mathcal{R}_0$ <u>is an open set then</u>

$$H_u(t,x_0,y_0,u_0,p) = 0.$$

This result follows from (4.18) and the relation of $F(t,x,y,u)$ to $H(t,x,y,u,p)$.

4.3 The Second Variation

In this section we suppose that $\mathcal{R}_0$ is an open set. Also suppose that $I_\gamma(y_0) = 0$, $\gamma = 1,\ldots,p'$. It is then true that $J_\rho(y_0) = 0$, $\rho = 1,\ldots,p+n$.

Let $\mu(t) = (\mu'(t),\ldots,\mu^q(t))$, $a \le t \le b$, be a PWC function on $[a,b]$ and set

$$u(t,\varepsilon) = u_0(t) + \varepsilon\mu(t).$$

In view of standard imbedding theorems, e.g. [5, Chap. IX, Sec. 3], applied successively to the intervals $[a,a+\tau]$, $[a+\tau,a+2\tau],\ldots$ the equations

$$\dot{y}^i = f^i(t,x,y,u), \quad y(t) = \alpha(t), \quad a-\tau \le t \le a$$

with $u = u(t,\varepsilon)$ have a one-parameter family of solutions

$$y(\varepsilon): \qquad y(t,\varepsilon),\ u(t,\varepsilon), \quad a \le t \le b$$

for $|\varepsilon| < \varepsilon_0$. Also $y(0) = y_0$. The functions $y(t,\varepsilon)$ are continuous on $[a,b]$ and have continuous derivatives with respect to ε. The arc

$$\eta: \qquad \eta(t) = \frac{\partial y}{\partial\varepsilon}(t,0), \quad \mu(t) = \frac{\partial u}{\partial\varepsilon}(t,0)$$

is called the variation of $y(\varepsilon)$ along y_0. The arc η satisfies the equations of variations for $i,j = 1,\ldots,n$; $k = 1,\ldots,q$:

$$\dot{\eta}^i(t) = f^i_{x^j}(t,x_0,y_0,u_0)\xi^j(t) + f^i_{y^j}(t,x_0,y_0,u_0)\eta^j(t) + f^i_{u^k}(t,x_0,y_0,u_0)\mu^k(t)$$

where $\xi^j(t) = \eta^j(t-\tau)$. A solution η of these equations will be called a <u>differentially admissible variation</u>.

Now consider the functionals $I_\gamma(y(\varepsilon))$, $\gamma = 0,1,\dots,p$, and $J_\rho(y(\varepsilon))$, $\rho = 0,1,\dots,p+n$, defined in Section 4.2. Upon differentiation with respect to ε at $\varepsilon = 0$ we see that

$$\frac{d}{d\varepsilon} I_\gamma(y(\varepsilon))\Big|_{\varepsilon=0} = \int_a^b [L_{\gamma x^i}\xi^i(t) + L_{\gamma y^i}\eta^i(t) + L_{\gamma u^k}\mu^k(t)]dt,$$

$$\frac{d}{d\varepsilon} J_\rho(y(\varepsilon))\Big|_{\varepsilon=0} = \int_a^b [f_{\rho x^i}\xi^i(t) + f_{\rho y^i}\eta^i(t) + f_{\rho u^k}\mu^k(t)]dt$$

with the partial derivatives of L_γ and f_ρ evaluated at $(t,x_0(t),y_0(t),u_0(t))$. Set $I'_\gamma(y_0,\eta) = \frac{d}{d\varepsilon} I_\gamma(y(\varepsilon))\Big|_{\varepsilon=0}$ and set $J'_\rho(y_0,\eta) = \frac{d}{d\varepsilon} J_\rho(y(\varepsilon))\Big|_{\varepsilon=0}$. After the linear change of variable $t = s+\tau$ in the terms involving $\xi^i(t)$, we find that

$$I'_\gamma(y_0,\eta) = \int_a^{b-\tau} \{[L_{\gamma y^i}(\ ) + L_{\gamma x^i}(+\tau)]\eta^i(t) + L_{\gamma u^k}(\ )\mu^k(t)\}dt$$

$$+ \int_{b-\tau}^{b} [L_{\gamma y^i}(\ )\eta^i(t) + L_{\gamma u^k}(\ )\mu^k(t)]dt,$$

and

$$J'_\rho(y_0,\eta) = \int_a^{b-\tau} \{[f_{\rho y^i}(\ ) + f_{\rho x^i}(+\tau)]\eta^i(t) + f_{\rho u^k}(\ )\mu^k(t)\}dt$$

$$+ \int_{b-\tau}^{b} [f_{\rho y^i}(\ )\eta^i(t) + f_{\rho u^k}(\ )\mu^k(t)]dt.$$

The arguments appearing in the empty parentheses are $(t,x_0(t),y_0(t),u_0(t))$ and those appearing in $(+\tau)$ are $(t+\tau,y_0(t),z_0(t),v_0(t))$ where $z_0(t) = y_0(t+\tau)$. Recall that the f_ρ were defined so that the coefficients of $\eta^i(t)$ in $J'_\rho(y_0,\eta)$ are zero. Hence

$$J_\rho'(y_0,\eta) = \int_a^b f_{\rho u^k}[t,x_0(t),y_0(t),u_0(t)]\mu^k(t)dt.$$

Note also that since

$$I_\gamma\big(y(\varepsilon)\big) = J_\gamma\big(y(\varepsilon)\big),\ J_{p+i}\big(y(\varepsilon)\big) = y^i(b,\varepsilon) - \beta^i,\ \gamma = 0,1,\ldots,p;\ i = 1,\ldots,n,$$

we have that

$$I_\gamma'(y_0,\eta) = J_\gamma'(y_0,\eta),\quad J_{p+i}'(y_0,\eta) = \eta^i(b), \qquad \gamma = 0,1,\ldots,p;\ i = 1,\ldots,n. \tag{4.22}$$

LEMMA 4.3.1 <u>Given a set of</u> N <u>differentially admissible variations</u>

η_σ: $\quad \eta_\sigma(t),\ \mu_\sigma(t),\quad a \le t \le b,\quad \sigma = 1,\ldots,N,$

<u>there is an</u> N-<u>parameter family</u>

$y(\varepsilon)$: $\quad y(t,\varepsilon_1,\ldots,\varepsilon_N),\ u(t,\varepsilon_1,\ldots,\varepsilon_N),\quad a \le t \le b$

<u>of arcs in</u> $\mathfrak{A}$ <u>containing</u> y_0 <u>for</u> $\varepsilon = 0$ <u>and having</u> η_σ <u>as its variations along</u> y_0 <u>in the sense that</u>

$$\eta_\sigma^i(t) = \frac{\partial y^i}{\partial \varepsilon_\sigma}(t,0),\ \mu_\sigma^k(t) = \frac{\partial u^k}{\partial \varepsilon_\sigma}(t,0),\quad i = 1,\ldots,n;\ k = 1,\ldots,q.$$

The proof of this lemma is given in several places for the case with no delayed arguments, e.g. [6, p.273]. The argument in the present case is the same, and so it will not be given here.

The arc y_0 will be said to be normal if there exists $p+n$ differentially admissible variations

η_σ: $\qquad \eta_\sigma(t),\ \mu_\sigma(t), \quad \sigma = 1,\ldots,p+n$

such that $\eta_\sigma(t) = 0$ for $a - \tau \leq t \leq a$ and such that the determinant

$$\begin{vmatrix} I'_\gamma(y_0,\eta_\sigma) \\ \eta^i_\sigma(b) \end{vmatrix} \qquad \begin{matrix} \gamma = 1,\ldots,p;\ i = 1,\ldots,n \\ \sigma = 1,\ldots,p+n \end{matrix}$$

is different from zero. In view of (4.22), this determinant may be put in the form

$$|J'_\rho(y_0,\eta_\sigma)| \qquad \rho,\sigma = 1,\ldots,p+n$$

The following two results may now be proved.

THEOREM 4.3.1 <u>Suppose that</u> y_0 <u>is normal and that</u>

η: $\qquad \eta(t),\ \mu(t)$

<u>is a differentially admissible variation having</u> $\eta(b) = 0,\ \eta(t) \equiv 0,$ $a - \tau \leq t \leq a,$ <u>and</u>

$$I'_\gamma(y_0,\eta) \leq 0, \quad 1 \leq \gamma \leq p'; \quad I'_\gamma(y_0,\eta) = 0, \quad p' < \gamma \leq p.$$

<u>There exists a one parameter family of differentially admissible arcs</u>

$y(\epsilon)$: $\qquad y(t,\epsilon),\ u(t,\epsilon), \quad |\epsilon| < \delta$

<u>containing</u> y_0 <u>for</u> $\epsilon = 0$, <u>joining the endpoints of</u> y_0, <u>satisfying the relations</u>

$$I_\gamma\big(y(\epsilon)\big) = \epsilon I'_\gamma(y_0,\eta),$$

<u>and having</u> η <u>as its variation along</u> y_0.

THEOREM 4.3.2 <u>If</u> y_0 <u>is normal, then the multiplier</u> λ_0 <u>is positive and hence it can be chosen as one.</u>

The same results for problems without delayed argument are to be found in [6, pp.274-75]. The proofs of Theorems 4.3.1 and 4.3.2 are word for word the same as in [6] and so will be omitted.

Recall that with $\lambda_0 = 1$,

(4.23) $H(t,x,y,u,p) = p_i f^i - L_0 - \lambda_\gamma L_\gamma, \quad i = 1,\ldots,n;\ \gamma = 1,\ldots,p.$

Also

(4.24) $F(t,x,y,u) = -H\big(t,x,y,u,p(t)\big) - \dot{p}_i(t)y^i, \quad i = 1,\ldots,n.$

The integral

(4.25) $$J(y) = \int_a^b F\big(t,x(t),y(t),u(t)\big)dt + c$$

has the property that

$$J(y) = I_0(y) + \lambda_\gamma I_\gamma(y)$$

for

$$c = p_i(b)y^i(b) - p_i(a)y^i(a).$$

The second variation of J along y_0 is given by the expression

$$J_2(y_0,\eta) = \int_a^b 2\omega(t,\xi,\eta,\mu)dt$$

where

$$2\omega(t,\xi,\eta,\mu) = F_{x^ix^j}\xi^i\xi^j + 2F_{x^iy^j}\xi^i\eta^j + F_{y^iy^j}\eta^i\eta^j$$

$$+ 2F_{y^iu^k}\eta^i\mu^k + 2F_{x^iu^k}\xi^i\mu^k + F_{u^ku^\ell}\mu^k\mu^\ell$$

$$= -H_{x^ix^j}\xi^i\xi^j - 2H_{x^iy^j}\xi^i\eta^j - H_{y^iy^j}\eta^i\eta^j - 2H_{y^iu^k}\eta^i\mu^k$$

$$-2H_{x^iu^k}\xi^i\mu^k - H_{u^ku^\ell}\mu^k\mu^\ell, \quad i,j = 1,\ldots,n; \; k,\ell = 1,\ldots,q.$$

The arguments of the partial derivatives of F are $(t,x_0(t),y_0(t),u_0(t))$ and those of H are $(t,x_0(t),y_0(t),u_0(t),p(t))$.

Let Γ be the class of all differentially admissible variations having $\eta(t) \equiv 0$, $a - \tau \le t \le a$, $\eta(b) = 0$ and satisfying the relations

$$I_\gamma'(y_0,\eta) \le 0, \quad 1 \le \gamma \le p'; \quad I_\gamma'(y_0,\eta) = 0, \quad p' < \gamma \le p.$$

Denote by Γ' the class of all variations η in Γ which have the further property that $I_\gamma'(y_0,\eta) = 0$ for all indices $\gamma \le p'$ for which $\lambda_\gamma > 0$.

For definiteness we state precisely the problem with which we are concerned. Minimize the functional

$$I_0(y) = \int_a^b L_0\big(t,x(t),y(t),u(t)\big)dt$$

on the class of all arcs

$$y: \qquad y(t),\; u(t)$$

satisfying the system of differential-difference equations

$$y^i = f^i(t,x,y,u),$$

the set of initial and terminal conditions

$$y^i(t) = \alpha^i(t), \quad a - \tau \le t \le a, \quad y^i(b) = \beta^i$$

and a set of isoperimetric relations

$$I_\gamma(y) \leq 0, \quad 1 \leq \gamma \leq p'$$

$$I_\gamma(y) = 0, \quad p' < \gamma \leq p.$$

The region $\mathcal{R}$ is the same as in section 3.1 and $\mathcal{R}_0$ is an open subset of $\mathcal{R}$. We further assume that at a minimizing arc y_0, $I_\gamma(y_0) = 0$, $1 \leq \gamma \leq p'$.

THEOREM 4.3.3 If y_0 is a normal minimizing arc for the above problem, the second variation $J_2(y_0,\eta)$ of J along y_0 has the property that

$$J_2(y_0,\eta) \geq 0$$

for all differentially admissible variations η belonging to class Γ'.

In order to prove this theorem let

$$\eta: \qquad \eta(t),\ \mu(t)$$

be a variation in Γ'. In view of Theorem 4.3.1, there exists a one-parameter family

$$y(\varepsilon): \qquad y(t,\varepsilon),\ u(t,\varepsilon), \qquad |\varepsilon| < \delta$$

of differentially admissible arcs joining the endpoints of y_0, containing y_0 for $\varepsilon = 0$, and having

$$I_\gamma\big(y(\varepsilon)\big) = \varepsilon I'_\gamma(y_0,\eta).$$

Since η is in Γ', we have

$$I_\gamma\big(y(\varepsilon)\big) \leq 0, \quad 1 \leq \gamma \leq p'; \quad I_\gamma\big(y(\varepsilon)\big) = 0, \quad p' < \gamma \leq p'$$

for ε on the range $0 \leq \varepsilon < \delta$.

If $\gamma \leq p'$ and $\lambda_\gamma > 0$, we then have $I'_\gamma(y_0, \eta) = 0$ so that the identity

$$\lambda_\gamma I_\gamma\big(y(\varepsilon)\big) = 0, \quad |\varepsilon| < \delta$$

is true. It follows from the definition (4.25) of $J(y)$ that

$$J\big(y(\varepsilon)\big) = I_0\big(y(\varepsilon)\big), \quad |\varepsilon| < \delta. \tag{4.26}$$

Using the relation (4.26) with $0 \leq \varepsilon < \delta$ and the minimizing property of y_0, the function

$$g(\varepsilon) = J\big(y(\varepsilon)\big) = I_0\big(y(\varepsilon)\big)$$

satisfies the inequality

$$g(\varepsilon) \geq g(0) = J(y_0) = I_0(y_0), \quad 0 \leq \varepsilon < \delta.$$

Now

$$g'(0) = \int_a^b [F_{x^i}(\ )\xi^i(t) + F_{y^i}(\ )\eta^i(t) + F_{u^k}(\ )\mu^k(t)]dt, \ i = 1,\dots,n;\ k = 1,\dots,q.$$

where the arguments appearing in the open parentheses are $\big(t, x_0(t), y_0(t), u_0(t)\big)$. The same linear change of variable as used previously yields

$$g'(0) = \int_a^{b-\tau} \{[F_{y^i}(\ ) + F_{x^i}(+\tau)]\eta^i(t) + F_{u^k}(\ )\mu^k(t)\}dt$$

$$+ \int_{b-\tau}^{b} [F_{y^i}(\ )\eta^i(t) + F_{u^k}(\ )\mu^k(t)]dt.$$

Here $(\ )$ means $\big(t, x_0(t), y_0(t), u_0(t)\big)$ and $(+\tau)$ means $\big(t+\tau, y_0(t), z_0(t), v_0(t)\big)$.

Using the definitions (4.24) of F and (4.23) of H, we see that

$$F_{y^i}(\ ) + F_{x^i}(+\tau) = -\mathring{p}_i(t) - H_{y^i}(\ ) - H_{x^i}(+\tau) = 0, \quad a \le t \le b - \tau$$

$$F_{y^i}(\ ) = -\mathring{p}_i(t) - H_{y^i}(\ ) = 0, \quad b - \tau \le t \le b$$

$$F_{u^j}(\ ) = -H_{u^j}(\ ) = 0 \quad j = 1,\ldots,q,$$

since y_0 is a minimizing arc. Therefore

$$g'(0) = 0.$$

Hence

$$g''(0) \geq 0.$$

However, note that $g''(0) = J''(y_0,\eta)$. Hence the theorem is proved.

4.4 Sufficient Conditions

In this section we modify the problem considered in Section 4.1 by assuming that the isoperimetric inequalities

$$I_\gamma(y) \le 0, \quad 1 \le \gamma \le p'$$

are indeed equalities, i.e.

$$I_\gamma(y) = 0, \quad 1 \le \gamma \le p'.$$

We also assume that $\mathcal{R}_0$ is an open set. Denote by $\mathcal{O}_0$ this new class of admissible functions.

The following theorem is quite similar to Theorem 3.2.1 for simple integral problems. Suppose that y_0 is a normal arc which satisfies the following conditions:

a. There exist multipliers $\lambda_0 = 1$, λ_γ, $p_i(t)$, $(\gamma = 1,\ldots,p;$ $i = 1,\ldots,n)$ such that y_0 with $p_i(t)$ and $\lambda_0,\ldots,\lambda_p$ satisfies the equations

$$\dot{p}_i = -H_{y^i}(t,x,y,u,p)-H_{x^i}\big(t+\tau,y,z,v,p(t+\tau)\big), \quad a \le t \le b-\tau,$$

$$\dot{p}_i = -H_{y^i}(t,x,y,u,p), \qquad b-\tau \le t \le b; \qquad i = 1,\ldots,n.$$

$$H_{u^k}(t,x,y,u,p) = 0, \qquad a \le t \le b; \qquad k = 1,\ldots,q.$$

b. $$H(t,x_0,y_0,u_0,p)-H(t,x,y,u,p)+(y^i-y_0^i)[H_{y^i}(t,x_0,y_0,u_0,p)$$
$$+H_{x^i}\big(t+\tau,y_0,z_0,v_0,p(t+\tau)\big)] \ge 0, \qquad a \le t \le b-\tau,$$

$$H(t,x_0,y_0,u_0,p)-H(t,x,y,u,p)+(y^i-y_0^i)H_{y^i}(t,x_0,y_0,u_0,p) \ge 0$$
$$b-\tau \le t \le b.$$

for all y in $\mathcal{O}_0$.

THEOREM 4.4.1 Under the above conditions, the arc y_0 furnishes I_0 with a global minimum.

In order to prove this result consider the following inequalities:

$$I_0(y) - I_0(y_0) = \int_a^b [L_0(t,x,y,u) - L_0(t,x_0,y_0,u_0)]dt$$

$$= \int_a^b \{p_i(t)[f^i(t,x,y,u)-f^i(t,x_0,y_0,u_0)]-[H(t,x,y,u,p)-H(t,x_0,y_0,u_0,p)]\}dt$$

$$= \int_a^b \{p_i(t)[\dot{y}^i(t)-\dot{y}_0^i(t)]-[H(t,x,y,u,p)-H(t,x_0,y_0,u_0,p)]\}dt.$$

An integration by parts yields the relation

$$\int_a^b \{-\dot{p}_i(t)[y^i(t)-y_0^i(t)]-[H(t,x,y,u,p)-H(t,x_0,y_0,u_0,p)]\}dt+p_i(t)[y^i(t)-y_0^i(t)]\Big|_a^b$$

Using condition a, we find that the preceding becomes

$$\int_a^b [H(t,x_0,y_0,u_0,p)-H(t,x,y,u,p)]dt$$

$$+ \int_a^{b-\tau} \{[y^i(t)-y_0^i(t)]H_{y^i}(t,x_0,y_0,u_0,p(t))+H_{x^i}(t+\tau,y_0,z_0,v_0,p(t+\tau))\}dt$$

$$+ \int_{b-\tau}^b [y^i(t)-y_0^i(t)]H_{y^i}(t,x_0,y_0,u_0,p(t))dt.$$

An application of condition b yields the result that $I_0(y) - I_0(y_0) \geq 0$ for all y in $\mathcal{O}\!\mathcal{l}$.

The above theorem is applicable to many examples with linear differential equations and quadratic cost functionals, e.g. [2, p.547].

One may of course prove a sufficient condition for the problem considered here by the indirect method of Hestenes. Such a theorem is quite like Theorem 2.1 of [9]. The proof of such a theorem is essentially to be found in [9] and also in the proof of our Theorem 3.3.1.

BIBLIOGRAPHY

1. Bliss, G. A. Lectures on the Calculus of Variations. Chicago: The University of Chicago Press, 1946.

2. Chyung, D. H. & E. B. Lee. "Linear Optimal Systems with Time Delay." SIAM Journal on Control, vol. 4 (1966), pp. 548-575.

3. El'sgol'c, L. E. Qualitative Methods of Mathematical Analysis. Providence, R. I.: American Mathematical Society, 1964.

4. Ewing, G. M. "Sufficient Conditions for Global Extrema in the Calculus of Variations." The Journal of the Astronautical Sciences, XII, pp. 102-105.

5. Graves, L. M. Theory of Functions of Real Variables. Second Edition. New York: McGraw-Hill Book Co., Inc., 1956.

6. Hestenes, M. R. Calculus of Variations and Optimal Control Theory. New York: John Wiley & Sons, Inc., 1956.

7. Hestenes, M. R. "On Variational Theory and Optimal Control Theory." SIAM Journal on Control, vol. 3 (1965), pp. 23-48.

8. Mookini, Edwin H. Necessary Conditions and Sufficient Conditions for an Optimal Control Problem in the Calculus of Variations. Dissertation, University of California, Los Angeles, 1964.

9. Mookini, Edwin H. "Sufficient Conditions for an Optimal Control Problem in the Calculus of Variations." Pacific Journal of Mathematics, vol. 17 (1966), pp. 263-278.

10. Oğuztöreli, M. N. Time-Lag Control Systems. New York: Academic Press, 1966.

11. Pontryagin, L. S. et. al. The Mathematical Theory of Optimal Processes. New York: Interscience Publishers, 1962.

12. Reid, W. T. "A Boundary Value Problem Associated with the Calculus of Variations." American Journal of Mathematics, vol. 54 (1932), pp. 769-790.

13. Reid, W. T. "Analogues of the Jacobi Condition for the Problem of Mayer in the Calculus of Variations." Annals of Mathematics, vol. 35 (1934), pp. 836-848.

14. Reid, W. T. "Sufficient Conditions by Expansion Methods for the Problem of Bolza in the Calculus of Variations." Annals of Mathematics, vol. 38 (1937), pp. 622-678.