NECESSARY CONDITIONS FOR SINGULAR EXTREMALS IN THE CALCULUS OF VARIATIONS

BY

GOH BEAN SAN

18 November, 1966

A thesis presented for the degree of
Doctor of Philosophy at the University of
Canterbury, Christchurch, New Zealand.

PREFACE

The original objective of this research was to derive necessary conditions which would enable us to determine the nature of the intermediate - thrust arcs of optimal rocket trajectories, which Lawden had obtained in closed form. From the calculus of variations point of view these arcs are known as singular extremals. Such extremals can satisfy the classical Clebsch-Legendre condition marginally (equality). In such a case, it is necessary to discover further inequality conditions by reference to which the nature of these extremals can be decided. Several authors have previously obtained such generalized Clebsch conditions, which are however limited to the variation of one control variable at the time. Here the complete generalized Clebsch condition is presented and this is stated as theorem 2 in chapter 4. In section 4.8 it is shown that the positive definiteness of this generalized Clebsch condition plays a role in the derivation of the singular extremals similar to that of the positive definiteness of the classical Clebsch condition in the derivation of regular extremals. The theory is applied mostly to problems from the theory of optimal control.

I wish to express my thanks to my research supervisor Professor D.F. Lawden for his many invaluable suggestions and in particular the original problem.

CONTENTS

CHAPTER 1. THE BOLZA PROBLEM

1.1. INTRODUCTION

The Calculus of Variations is in a sense an extension of the Differential Calculus. It is mainly concerned with the theory of maxima and minima. However, the quantities to be extremized belong to a class of functionals, i.e. functions of functions. One of the most important problems of the Calculus of Variations is the Bolza problem, which was first formulated in 1913. Theoretically this problem is equivalent to the Lagrange problem which was formulated in 1770 and the Mayer problem which was formulated in 1878. An interesting short history of these problems can be found in Bliss [1].

In the nomenclature of Bliss [2], the Bolza problem can be formulated thus: $y_i(x)$, $i = 1,2,\ldots,n$ is a set of functions defined over an interval $a < x < b$ and satisfying differential constraints

$$\phi^{\beta}(x,y,y') = 0, \ \beta = 1,2,\ldots\ldots,m < n, \qquad (1.1.1)$$

where y,y' denote the whole set of functions and their derivatives. The derivatives $y_i'(x)$ are assumed to be continuous except for a finite number of finite discontinuities. x_1,x_2 are endpoints satisfying $a < x_1 < x_2 < b$. The values of the functions $y_i(x)$ at these endpoints are required to satisfy end conditions

$$\psi^{\mu}[x_1, y(x_1), x_2, y(x_2)] = 0, \quad \mu = 1,2,\ldots p \leq 2n+2. \quad (1.1.2)$$

The problem is to determine the set of functions (if it exists) and endpoints satisfying these end conditions and the differential constraints, which minimizes a functional J given by

$$J \equiv g[x_1, y(x_1), x_2, y(x_2)] + \int_{x_1}^{x_2} f(x,y,y')dx, \qquad (1.1.3)$$

where g and f are known functions.

1.2. NOTATIONS AND ASSUMPTIONS

The notation which will be employed is similar to that used by Bliss in [2]. Thus partial derivatives will be indicated by subscripts. The repeated indices summation convention will be used unless otherwise stated. We say that $f(x,y,z) \in C^n$ in an open connected region Γ of the (x,y,z)-space, if all partial derivatives of order n exist and are continuous. Again we say that $y(x) \in D^n$ in an open connected subset Γ of the Euclidean line if $y(x) \in C^{n-1}$ in Γ and the nth derivative $y^n(x)$ is continuous except for a finite number of finite discontinuities at which left- and right-hand limits exist.

Let Γ_1 be an open region of $(2n+1)$-dimensional points (x,y,y') in which $\phi^{\beta}(x,y,y') \in C^3$ and $f(x,y,y') \in C^3$. Then the matrix

$$\left(\phi^{\beta}_{y'_i}\right) \tag{1.2.1}$$

will be assumed to have rank m everywhere in Γ_1.

Similarly let Γ_2 be an open region of $(2n+2)$-dimensional points $(x_1, y_{i1}, x_2, y_{i2})$ in which $\psi^{\mu}(x_1, y_{i1}, x_2, y_{i2}) \in C^3$ and $g(x_1, y_{i1}, x_2, y_{i2}) \in C^3$. Then the matrix

$$(\psi^{\mu}_1 \ \psi^{\mu}_{i1} \ \psi^{\mu}_2 \ \psi^{\mu}_{i2}) \tag{1.2.2}$$

will be assumed to be of rank p in Γ_2. Here

$$\psi^{\mu}_1 = \frac{\partial\psi^{\mu}}{\partial x_1} \text{ and } \psi^{\mu}_{i1} = \frac{\partial\psi^{\mu}}{\partial y_{i1}} , \tag{1.2.3}$$

with similar notations for other derivatives.

Given functions $y_i(x)$ define an arc E in an $(n+1)$ dimensional space in which $(x, y_1, y_2, \ldots, y_n)$ are coordinates. An arc E is said to be <u>admissible</u> if its elements (x, y, y') and $(x_1, y_{i1}, x_2, y_{i2})$ belong to the regions Γ_1 and Γ_2 defined above and furthermore satisfy the differential constraints (1.1.2) and the end conditions (1.1.3). This definition of admissibility is more restrictive than that of Bliss [2] because the arc E is required to satisfy the differential constraints <u>and</u> the end conditions.

1.3. <u>THE CLASSICAL NECESSARY CONDITIONS</u>

The standard approach in the Calculus of Variations is to assume that a certain admissible arc E with elements (x,y,y') and end conditions $(x_1,y_{i_1},x_2,y_{i_2})$ is the minimizing arc. Four sets of necessary conditions are then deduced which must be satisfied by <u>any</u> minimizing arc. Normally, it is expected that these four sets of necessary conditions provide sufficient information to characterize uniquely the minimizing arc. These four sets of necessary conditions are known as the Multiplier Rule, the Weierstrass condition, the Clebsch (Legendre) condition and the Jacobi condition.

<u>The Multiplier Rule</u>: For a minimizing arc E, it is necessary that there exists a non-negative constant λ_0 and constants e^μ, c^i together with functions $\lambda_\beta(x)$ such that the function

$$F(x,y,y',\lambda) \equiv \lambda_0 f(x,y,y') + \lambda_\beta \phi^\beta(x,y,y') \qquad (1.3.1)$$

satisfies the du Bois-Reymond equations

$$F_{y_i'}(x,y,y',\lambda) = \int_{x_1}^{x} F_{y_i}(x,y,y',\lambda)dx + c^i \qquad (1.3.2)$$

along arc E, and furthermore, the Transversality condition

$$[(F - y_i'F_{y_i'})dx + F_{y_i'}dy_i]_1^2 + \lambda_0 dg + e^\mu d\psi^\mu \equiv 0 \quad (1.3.3)$$

for all arbitrary differentials dx_1, dy_{i_1}, dx_2 and dy_{i_2}, must be satisfied at the ends of the arc E.

The first important consequence of the Multiplier Rule is that at each point of the arc E which is not a corner (i.e. $y_i'(x)$ are continuous) the du Bois - Reymond equations can be differentiated and this leads to the Euler-Lagrange equations

$$\frac{d}{dx} F_{y_i'} = F_{y_i} \ . \qquad (1.3.4)$$

The second set of important consequences are the Weierstrass-Erdmann corner conditions which state that the functions $F_{y_i'}$ and $(F - y_i' F_{y_i'})$ have at a corner well defined right and left limits which are equal.

Definition: A smooth arc satisfying the Euler-Lagrange equations is known as an extremal.

Normality Assumptions: Hereafter it will be assumed that the minimizing arc E is normal, i.e. it satisfies the Multiplier Rule with a unique set of multipliers of the form $\lambda_0 = 1$ and $\lambda_\beta(x)$ (see Bliss [2, p.214]).

The Weierstrass Condition: Each element (x,y,y') belonging to a normal minimizing arc E, satisfying the Multiplier Rule, must satisfy the inequality

$$E(x,y,y',Y',\lambda) \geq 0 \qquad (1.3.5)$$

for all admissible sets (x,y,Y') where

$$E(x,y,y',Y',\lambda) \equiv F(x,y,Y',\lambda) - F(x,y,y',\lambda) - (Y_i' - y_i')F_{y_i'}(x,y,y',\lambda). \qquad (1.3.6)$$

The Clebsch Condition: Each element (x,y,y') belonging to a normal minimizing arc E, satisfying the Multiplier Rule, must satisfy the inequality

$$F_{y_i' y_j'} \pi_i \pi_j \geqslant 0, \tag{1.3.7}$$

for all sets of numbers π_i satisfying the equations

$$\phi^{\beta}_{y_i'}(x,y,y',\lambda)\pi_i = 0. \tag{1.3.8}$$

For a Bolza problem without any differential constraint, this condition is better known as the Legendre condition because, in 1786, Legendre obtained such a condition for the simplest problem of the Calculus of Variations by means of the second variation. The Clebsch condition in the above form was first proved for the Lagrange problem by Clebsch in 1858 by means of a transformation of the second variation. The Clebsch condition deduced in this manner is limited in applicability because the transformation is valid only at points of the minimizing arc for which the determinant

$$\Delta \equiv \begin{vmatrix} F_{y_i' y_j'} & \phi^{\gamma}_{y_i'} \\ \phi^{\beta}_{y_j'} & 0 \end{vmatrix} \tag{1.3.9}$$

does not vanish. By means of continuity arguments, it is then possible to extend the applicability of this Clebsch condition to minimizing arcs along which Δ vanishes at only a finite number of points. However, the Clebsch condition can also be deduced as a direct consequence of the Weierstrass condition, Bliss [2, p.224], and is then applicable even to a minimizing arc along which the determinant Δ vanishes identically.

The Jacobi Condition: Let E be a normal non-singular extremal satisfying the Multiplier Rule and the Clebsch condition. Then the second variation of the functional J along arc E must be non-negative for all admissible variations.

Arc E is non-singular if the determinant Δ displayed in (1.3.9) does not vanish along E. A set of admissible variations along E, is a set of two parameters ξ_1, ξ_2 and n functions $\eta_i(x)$ satisfying the equations of variation

$$\Phi^\beta \equiv \frac{\partial\phi^\beta}{\partial y_i'}\eta_i' + \frac{\partial\phi^\beta}{\partial y_i}\eta_i = 0, \qquad (1.3.10)$$

$$\Psi^\mu \equiv \left(\frac{\partial\psi^\mu}{\partial x_1} + y_{i1}' \frac{\partial\psi^\mu}{\partial y_{i1}}\right)\xi_1 + \frac{\partial\psi^\mu}{\partial y_{i1}}\eta_{i1} +$$

$$\left(\frac{\partial\psi^\mu}{\partial x_2} + y_{i2}' \frac{\partial\psi^\mu}{\partial y_{i2}}\right)\xi_2 + \frac{\partial\psi^\mu}{\partial y_{i2}}\eta_{i2} = 0, \qquad (1.3.11)$$

where $y_{i1} = y_i(x_1)$, $y_{i2} = y_i(x_2)$ and the partial derivatives in (1.3.10) and (1.3.11) are to be calculated along E. It is demonstrated by Bliss [2] that for a normal arc E, variations satisfying the equations of variation (1.3.10) and (1.3.11) certainly exist.

Given such a set of variations along an extremal arc E for which the second order derivatives $y_i''(x)$ are continuous, the second variation J, is given by

$$J_2 \equiv 2\gamma[\xi_1, \eta(x_1), \xi_2, \eta(x_2)] + \int_{x_1}^{x_2} 2\omega(x, \eta, \eta')dx, \qquad (1.3.12)$$

where 2γ is a certain homogenous quadratic form in its arguments and

$$2\omega \equiv R_{ij}\eta_i'\eta_j' + 2Q_{ij}\eta_i'\eta_j + P_{ij}\eta_i\eta_j . \qquad (1.3.13)$$

The coefficients R_{ij}, Q_{ij}, P_{ij} are evaluated along the arc E from the equations

$$R_{ij} = \frac{\partial^2 F}{\partial y_i' \partial y_j'}, \; Q_{ij} = \frac{\partial^2 F}{\partial y_i' \partial y_j}, \; P_{ij} = \frac{\partial^2 F}{\partial y_i \partial y_j}, \qquad (1.3.14)$$

where

$$F \equiv f + \lambda_\beta \phi^\beta, \qquad (1.3.15)$$

the $\lambda_\beta(x)$ being multipliers occurring in the Multiplier Rule.

1.4. THE ACCESSORY MINIMUM PROBLEM.

Historically the Clebsch condition and the Jacobi condition were first deduced from the second variation by means of the so called transformation theory which flourished into an elaborate theory in the second half of the nineteenth century. A summary of these transformations can be found in Bolza [3]. The most important of these transformations is that due to Clebsch, a version of which can be found in Reid [4]. The transformation theory of the second variation has been largely superseded by other methods especially by the accessory minimum problem approach, which was first proposed by Bliss [5] and extended by Smith [6]. The background to the introduction of this accessory minimum problem can be found in an interesting paper by Bliss [7].

The accessory minimum problem arises in the following manner. Let E be a normal extremal along which the second order derivatives $y_i''(x)$ are continuous and which satisfies the Multiplier Rule. Then the second variation J_2, displayed in (1.3.12) is well-defined irrespective of whether or not the determinant Δ of (1.3.9) is singular along E. Furthermore, without first requiring that the Clebsch condition be satisfied along E, the second variation $J_2(\xi,\eta)$ must always be non-negative for all admissible variations i.e.

$$J_2(\xi,\eta) \geqslant 0. \tag{1.4.1}$$

Hence any set of admissible variations $\xi_1,\xi_2,\eta_i(x)$ for which the second variation is zero, must minimize J_2 or else, there exists other sets of admissible variations, for which J_2 is negative. Thus, we are led to an auxiliary minimum problem which is similar to the standard Bolza problem (see equations (1.1.1) to (1.1.3)), except for the appearance of the parameters ξ_1,ξ_2. By a simple transformation, Bliss [2, p.232], this minimum problem can be transformed into the standard Bolza problem. However, this is unnecessary and instead, we shall use the modified forms of the necessary conditions when applying them to accessory extremals.

For the accessory minimum problem let the function corresponding to F of the Multiplier Rule be

$$\Omega(x,\eta,\eta',\ell) \equiv 2\omega(x,\eta,\eta') + \ell_\beta\left(\frac{\partial\phi^\beta}{\partial y_i'}\eta_i' + \frac{\partial\phi^\beta}{\partial y_i}\eta_i\right). \tag{1.4.2}$$

A particular set of admissible variations satisfying the conditions (1.3.10) and (1.3.11) is given by

$$\xi_1 = 0 = \xi_2, \ \eta_i(x) \equiv 0 \tag{1.4.3}$$

and for this set the second variation J_2 vanishes. The Multiplier Rule, which requires that there exists constants c_i and ϵ_μ and multipliers $\ell_\beta(x)$ such that

$$\Omega_{\eta_i'} = \int_{x_1}^{x} \Omega_{\eta_i} dx + c_i \tag{1.4.4}$$

and $$2d\gamma + \epsilon_\mu d\Psi^\mu + [(\Omega - \eta_i' \Omega_{\eta_i'})dx + \Omega_{\eta_i'} d\eta_i]_{x_1}^{x_2} \equiv 0, \tag{1.4.5}$$

for arbitrary differentials $d\xi_1, d\xi_2, d\eta_i(x_1)$ and $d\eta_i(x_2)$, is satisfied with the unique set of multipliers $\epsilon_\mu = 0$, $c_i = 0$ and $\ell_\beta(x) \equiv 0$. The uniqueness of this set of multipliers is a consequence of the assumption that arc E is normal.

Applying the Weierstrass condition to this accessory extremal we deduce that

$$\begin{aligned} E(x,0,0,\eta_i'^*,0) &\equiv \Omega(x,0,\eta_i'^*,0) \\ - \Omega(x,0,0,0) \quad &- (\eta_i'^* - \eta_i')\Omega_{\eta_i'}(x,0,0,0) \\ &= 2F_{y_i' y_j'} \eta_i'^* \eta_j'^* \geqslant 0 \end{aligned} \tag{1.4.6}$$

for all admissible sets $(x,0,\eta_i'^*)$ satisfying

$$\phi^\beta_{y_i'} \eta_i'^* = 0. \tag{1.4.7}$$

The inequaltiy condition (1.4.6) subjected to condition

(1.4.7) is the Clebsch condition for extremal arc E in the primary (original) Bolza problem. This is in fact one method of proving the Clebsch condition which is general enough to be applicable to extremals, along which the determinant Δ of (1.3.9) may vanish identically. This proof of the Clebsch condition can be found in Smith [6].

On the other hand if the Clebsch condition is considered as a consequence of the Weierstrass condition, (e.g. Bliss [2, p.224]) it would then be of some interest to examine what happens when the Clebsch condition is applied to the accessory extremal displayed in (1.4.3). As

$$\frac{\partial^2 \eta\Omega}{\partial\eta_i'\partial\eta_j'} = 2R_{ij} = 2\frac{\partial^2 F}{\partial y_i'\partial y_j'} \tag{1.4.8}$$

and

$$\frac{\partial\Phi^\beta}{\partial\eta_i'} = \frac{\partial\phi^\beta}{\partial y_i'}\ , \tag{1.4.9}$$

we are again led to the Clebsch condition in the primary Bolza problem.

The second and more important use of the accessory minimum problem is to test the Jacobi condition. This is discussed in Bliss [2]. Analytically this is a very difficult test to carry out.

CHAPTER 2. SINGULAR EXTREMALS

Definition: A minimizing arc E is said to be singular if the determinant Δ of (1.3.9) vanishes at any point on it. We shall study only the case where Δ vanishes at every point of E. The case where Δ vanishes at only a finite number of points of E has been examined by Morse and Leighton [8].

2.1. HISTORICAL BACKGROUND

Little research was done on singular extremals until the last few years. In Bolza [9,p.29], the case when the Euler equation degenerates into an algebraic identity, is examined. The functional to be minimized is then independent of the path of integration. Hereafter, such degenerate problems will not be examined.

The more interesting type of singular extremals appears to have been first examined in detail, by Mancill [10], who studied the minimization of the Linear Integral in the plane. The first and higher order variations were examined and necessary conditions were deduced. With the help of the Green's theorem on line integrals, the strengthened form of these necessary conditions provide a set of sufficient conditions for a strong proper relative minimum. This Green's theorem technique was independently developed and popularized by Miele [11], who used it extensively to study optimum flight paths. Recently, Haynes [12] has generalized the Green's theorem technique to higher dimensional singular problems of the calculus of variations.

Since 1950 there has been a tremendous interest in the applications of the calculus of variations to the study of optimal rocket trajectories and optimal control, see Lawden [13], Leitmann [14] and Paiewonsky [15]. Singular extremals were found to occur in several of these problems for which the Green's theorem technique was not applicable. Two major difficulties were encountered. Firstly, there was no general procedure by which the equations of the singular extremals could be derived. Secondly, the classical necessary conditions were found to be inadequate to decide the optimality of some of these singular extremals.

Important contributions towards the solution of the first difficulty were made by Lawden [13], Kelley [16] and Johnson [17]. In the theory of optimal rocket trajectories, Lawden derived the equations of the intermediate-thrust arcs which are now known as the Lawden's spirals and which are singular extremals. Kelley laid down a general transformation procedure which is unfortunately limited by the requirement that "closed form" solution of a system of non linear differential equations is required for the synthesis of the transformation. Johnson studied in great detail a class of singular extremals and laid down a rule for the derivation of these singular extremals. This rule involves the successive differentiations, with respect to the independent variable, the condition of singularity.

With respect to the second difficulty a breakthrough was made by Kelley [18] when he deduced a new necessary condition for a class of singular extremals. However, this necessary condition proved to be ineffective when applied to the Lawden's spirals. Several authors, Kopp and Moyer [19], Robbins [20], Gurley [21], and Tait [22], independently generalized Kelley's method and proved that the Lawden's spirals are non-optimal in the time-open case. Independently of these authors and by a different approach this writer Goh [23] arrived at the same conclusions concerning the Lawden's spirals. The method of this writer has been elaborated in Goh [24] and new results have been obtained for singular extremals involving multiple control variables.

2.2. MINIMIZATION OF LINEAR INTEGRALS IN THE PLANE

Consider the minimization of the functional

$$J \equiv \int_{x_1}^{x_2} [P(x,y) + Q(x,y)y']dx \ , \tag{2.2.1}$$

in the class of admissible arcs joining two fixed points and lying in the interior of a region Γ of the (x,y,y') space and where $P(x,y)$, $Q(x,y)$ are known functions of x and y. There are no differential constraints so that

$$\Delta = R = 0 \tag{2.2.2}$$

identically: the problem is accordingly singular.

The Euler-Lagrange equation for this problem is

$$\frac{d}{dx} Q = \frac{\partial P}{\partial y} + y' \frac{\partial Q}{\partial y} , \qquad (2.2.3)$$

$$\Longrightarrow \qquad \alpha(x,y) \equiv P_y - Q_x = 0 . \qquad (2.2.4)$$

This is an algebraic equation which in general, denotes a finite number of curves in the plane. Assuming that there exists an extremal satisfying this equation and also the given end conditions, it remains to be decided whether this extremal does, in fact, minimize J.

The Weierstrass condition requires that

$$y'^{*}Q - y'Q - (y'^{*} - y')Q \geqslant 0 , \qquad (2.2.5)$$

which is trivially satisfied. The Legendre (Clebsch) condition requires that

$$\frac{\partial^2}{\partial y'^2} (P + y'Q) \geqslant 0 , \qquad (2.2.6)$$

which is also trivially satisfied.

The second variation for an admissible variation $\eta(x)$ with $\eta(x_1) = 0 = \eta(x_2)$ is given by

$$J_2 = \int_{x_1}^{x_2} \left[2 \frac{\partial Q}{\partial y} \eta\eta' + \left(\frac{\partial^2 P}{\partial y^2} + y' \frac{\partial^2 Q}{\partial y^2} \right) \eta^2 \right] dx . \qquad (2.2.7)$$

Application of the Legendre condition to the accessory extremal $\eta(x) \equiv 0$, yields no further information as expected. However, integrating the first term in the integrand (22.7) by parts and employing the condition $\eta(x_1) = 0 = \eta(x_2)$, J_2 can be put in the form

$$J_2 = \int_{x_1}^{x_2} \left[\frac{\partial^2 P}{\partial y^2} - \frac{\partial^2 Q}{\partial y \partial x} \right] \eta^2 dx \ , \tag{2.2.8}$$

$$= \int_{x_1}^{x_2} \frac{\partial \alpha}{\partial y} \ \eta^2 dx \ , \tag{2.2.9}$$

where α is given by (2.2.4). Application of the Legendre condition to this new form of J_2 still yields no new condition. However, if we put

$$\eta(x) = \zeta'(x) \ , \tag{2.2.10}$$

so that J_2 takes the form

$$J_2 = \int_{x_1}^{x_2} \frac{\partial \alpha}{\partial y} \ \zeta'^2 dx \tag{2.2.11}$$

and now consider the accessory minimum problem with regard to the unknown function $\zeta(x)$ it is easily seen that the Legendre condition requires that

$$\frac{\partial \alpha}{\partial y} \geqslant 0 \ . \tag{2.2.12}$$

This could be considered as the generalized Legendre condition and has already been obtained by Mancill [10], who deduced it from the non-negativeness of the second variation in the form (2.2.8).

It should be noted that the end conditions $\eta(x_1) = 0 = \eta(x_2)$ can be disregarded after J_2 has been expressed in the form (2.2.9), since any function $\eta(x)$ not satisfying these conditions can be replaced by a modified function $\bar{\eta}(x)$, differing from $\eta(x)$ only in arbitrary small neighbourhoods of the endpoints and such that $\bar{\eta}(x_1) = 0 = \bar{\eta}(x_2)$; such a replacement will result in a change in J_2 which is also arbitrarily small and hence can be disregarded for the purpose of the condition $J_2 \geqslant 0$. It follows that $\zeta(x)$ is not required to satisfy end conditions on its derivative.

The transformation (2.2.10) is in fact a well known device, see Miele [25,p.110], Hestenes [26,p.74] and Berkovitz [27], which can be used to transform variational problems in which the derivative of the dependent variable does not occur, into the standard form of the Bolza problem. However, there is a valid objection to the use of this device here, arising from the fact that in the standard proof of the Weierstrass condition and hence the Legendre condition, it is assumed that the class $\zeta(x) \in D'$ whereas $\zeta'(x) = \eta(x) \in C^0$ and hence $\zeta(x) \in C'$. This objection may be removed by means of a "rounding argument", see Pars [28,p.10], which extends the applicability of the Weierstrass and Legendre conditions to variational problems for which the class of admissible curves belong to class C' instead of D'.

If the inequality (2.2.12) is satisfied marginally (equality) along an extremal, Mancill [10] has shown that the higher order variations yield the necessary conditions

$$\frac{\partial^k \alpha}{\partial y^k} = 0, \quad k = 2,3,\ldots,2n-2 \ , \qquad (2.2.13)$$

and

$$\frac{\partial^{2n-1} \alpha}{\partial y^{2n-1}} \geqslant 0 \ , \qquad (2.2.14)$$

assuming that the functions P(x,y) and Q(x,y) belong to the class C^{2n} in the region Γ. Note that $\partial\alpha/\partial y = 0$ implies $\partial\alpha/\partial x = 0$ and that (2.2.13) implies

$$\frac{\partial^k \alpha}{\partial x^k} = 0 \qquad (2.2.15)$$

and

$$\frac{\partial^k \alpha}{\partial y^\ell \partial x^m} = 0 \ , \qquad (2.2.16)$$

for all ℓ,m such that $\ell + m = k$. For differentiating $\alpha = 0$ with respect to x, it is easily seen that $\partial\alpha/\partial y = 0$ implies $\partial\alpha/\partial x = 0$. Then differentiating $\partial\alpha/\partial y = 0$ and $\partial\alpha/\partial x = 0$ with respect to x the above conditions are proved for k = 2 and in this manner these conditions can be proved for all admissible values of k.

Let us now consider the problem of the derivation of the equations of the singular extremal. By the implicit function theorem, equation (2.2.4) can be solved for y(x) if $\partial\alpha/\partial y \neq 0$. Thus if equality (2.2.12) is satisfied in the strengthened form, (2.2.4) can be used to derive the equation of the extremal.

Alternatively differentiating (2.2.4) with respect to x we obtain

$$(\partial\alpha/\partial y)y' + (\partial\alpha/\partial x) = 0 \ . \qquad (2.2.17)$$

If $\partial\alpha/\partial y$ does not vanish this differential equation has a one parameter general solution. The parameter is determined by subsituting the solution into (2.2.4) at a point $x^* \in [x_1, x_2]$. If this parameter is not uniquely determined at a certain point x^* a different value of x^* should be chosen.

If $\partial\alpha/\partial y$ vanishes identically for a certain set (x,y) then we have shown that $\partial\alpha/\partial x$ must also vanish unless the extremal is a straight line parallel to the y-axis in which case the variable y should be considered as the independent variable. Suppose that for a certain set (x,y), $\alpha(x)$, $\partial\alpha/\partial y$ and $\partial\alpha/\partial x$ vanish. Then the third order variation implies that $\partial^2\alpha/\partial y^2$ must vanish and hence $\partial^2\alpha/\partial x\partial y$ and $\partial^2\alpha/\partial x^2$ must also vanish, from (2.2.15) and (2.2.16).

Again assuming that α, $\partial\alpha/\partial y$, $\partial\alpha/\partial x$, $\partial^2\alpha/\partial y^2$, $\partial^2\alpha/\partial x\partial y$ and $\partial^2\alpha/\partial x^2$ all vanish for a certain set (x,y) the differentiation of $\partial^2\alpha/\partial y^2 = 0$ leads to

$$(\partial^3\alpha/\partial y^3)y' + (\partial^3\alpha/\partial y^2\partial x) = 0 \qquad (2.2.18)$$

and if $\partial^3\alpha/\partial y^3$ does not vanish this first order differential equation has a one parameter family general solution. This parameter must then be consisting determined from the vanishing of α and its first and second order partial derivatives at any point $x^* \in [x_1, x_2]$.

Finally it is obvious from this discussion that the quantities $\partial^{2p-1}\alpha/\partial y^{2p-1}$ for $p = 1,2,\ldots,n$ now play a role similar to $f_{y'y'}$ in the Hilbert's differentiability condition.

Example 1. Let $P(x,y) = y^4$, $Q(x,y) = xy^3$ and let the fixed end points be (0,0) and (1,0). Then

$$\alpha(x,y) = 3y^3 \tag{2.2.19}$$

$$\partial\alpha/\partial y = 9y^2 \tag{2.2.20}$$

$$\partial^2\alpha/\partial y^2 = 18y \tag{2.2.21}$$

$$\partial^3\alpha/\partial y^3 = 18 \tag{2.2.22}$$

$$> 0 \; . \tag{2.2.23}$$

The singular minimizing arc is given by

$$y(x) \equiv 0 \tag{2.2.24}$$

and with the help of the Green's Theorem on line integral, this extremal may be shown to give a proper absolute minimum. Note that $\alpha(x,y)$ is positive above the x-axis and negative below the x-axis.

CHAPTER 3. SINGULAR EXTREMALS IN THE (n+1)-DIMENSIONAL SPACE

3.1. A LEMMA

The following lemma will be of service:

Lemma 1. Let $R(x)$ be an $n \times n$ order symmetric matrix whose elements are well-defined real functions of x in $[x_1, x_2]$. Let R be of the form

$$R = \begin{pmatrix} R_1 & R_3^T \\ R_3 & 0 \end{pmatrix}, \tag{3.1.1}$$

where R_1 is a $m \times m$ order $(1 \leqslant m < n)$ matrix. If R is positive semidefinite in the interval $[x_1, x_2]$, then

$$R_3 \equiv 0. \tag{3.1.2}$$

Proof: At a point $x \in [x_1, x_2]$, consider a typical 2×2 order determinant

$$\Delta_2 = \begin{vmatrix} R_{1ii} & R_{3ri} \\ R_{3ri} & 0 \end{vmatrix}, \text{ (i,r not summed)}, \tag{3.1.3}$$

where R_{1ii} is an element from R_1 and R_{3ri} is an element of R_3. As matrix R is positive semidefinite hence

$$\begin{aligned} \Delta_2 &= R_{1ii} \cdot 0 - (R_{3ri})^2 \\ &\geqslant 0, \end{aligned} \tag{3.1.4}$$

$$\Rightarrow R_{3ri} = 0. \tag{3.1.5}$$

This is true for all admissible values of i and r and hence the matrix

$$R_3 = 0 .$$

As x is an arbitrary point of $[x_1,x_2]$, the matrix R_3 must be identically equal to the zero matrix. Q.E.D.

3.2. PRELIMINARY REMARKS

Consider the minimization of the functional

$$J \equiv g[x_1,y(x_1),x_2,y(x_2)] + \int_{x_1}^{x_2} f(x,y,y')dx, \qquad (3.2.1)$$

in the class of arcs $y_i(x)$, $i = 1,2,\ldots,n$, whose endpoints satisfy the equations

$$\psi^{\mu}[x_1,y(x_1),x_2,y(x_2)] = 0, \ \mu = 1,2,\ldots,p \leqslant 2n+2. \qquad (3.2.2)$$

This variational problem is in the form of the standard Bolza problem but for the absence of differential constraints. The analytical assumptions of section 1.2. will be assumed to be valid for this problem.

Along an extremal arc E, satisfying the Multiplier Rule and along which $y_i''(x)$ are assumed to be continuous, the second variation is well defined and is of the form

$$J_2 \equiv 2\gamma[\xi_1,\eta_i(x_1),\xi_2,\eta_i(x_2)] + \int_{x_1}^{x_2} 2\omega(x,\eta,\eta')dx, \qquad (3.2.3)$$

where 2γ is a certain quadratic form in its arguments and where

$$2\omega(x,\eta,\eta') \equiv \eta'^{T}R\eta' + 2\eta'^{T}Q\eta + \eta^{T}P\eta \qquad (3.2.4)$$

and normally $\qquad R = (\partial^2 f/\partial y_i'\partial y_j'), \qquad (3.2.5)$

$$Q = (\partial^2 f/\partial y_i' \partial y_j), \tag{3.2.6}$$

$$P = (\partial^2 f/\partial y_i \partial y_j). \tag{3.2.7}$$

It is quite possible that the matrices R, Q, P are not given by (3.2.5) to (3.2.7) as rearrangements may have been made on the quadratic form 2ω. Finally any set of admissible variations ξ_1, ξ_2, $\eta_i(x)$ must satisfy the equations of variation

$$\Psi^{\mu}\ [\xi_1, \eta_i(x_1), \xi_2, \eta_i(x_2)] = 0, \tag{3.2.8}$$

which are certain linear forms in their arguments.

Along a singular extremal

$$\Delta(x) = |R| \tag{3.2.9}$$

$$\equiv 0\ . \tag{3.2.10}$$

It will now be shown that if the classical Legendre condition is satisfied, then without any loss of generality, the matrix R, along a singular extremal, may be assumed to be of the form

$$R(x) \equiv \begin{pmatrix} R_1 & 0 \\ 0 & 0 \end{pmatrix}\ , \tag{3.2.11}$$

where R_1 is a $m \times m$ order matrix $(m < n)$ and where partition lines run between the mth and (m+1)th rows/columns.

Firstly, an important class of singular extremals consists of problems where the integrands $f(x,y,y')$ are expressions linear in $y'_{m+1}, y'_{m+2}, \ldots\ldots, y'_n$. Employing the indices $r,s = 1,2,\ldots,m$; $\rho,\nu = m+1, m+2,\ldots,n$, the matrix R is of the form

$$R = \begin{pmatrix} \dfrac{\partial^2 f}{\partial y'_r \partial y'_s} & \dfrac{\partial^2 f}{\partial y'_r \partial y'_\nu} \\ \dfrac{\partial^2 f}{\partial y'_\rho \partial y'_s} & 0 \end{pmatrix} = \begin{pmatrix} R_1 & R_3^T \\ R_3 & 0 \end{pmatrix}, \qquad (3.2.12)$$

where partition lines run between the mth and (m+1)th rows/columns. As the classical Legendre condition requires that R must be positive semidefinite, by Lemma 1 we are led to conclude that R_3 must vanish identically and hence R is of the form (3.2.11).

In the general case we know from matrix theory that, as R is symmetric and real there exists a nonsingular matrix such that $V^T RV$ is of the form (3.2.11). Using such a matrix V consider the nonsingular linear transformation

$$\eta = V\zeta. \qquad (3.2.13)$$

Differentiating,

$$\eta' = V\zeta' + V'\zeta \qquad (3.2.14)$$

and the last two equations lead to

$$\eta'^T R\eta' = \zeta'^T V^T RV\zeta' + 2\zeta'^T V^T RV'\zeta + \zeta^T V'^T RV'\zeta, \qquad (3.2.15)$$

$$2\eta'^T Q\eta = 2\zeta'^T V^T QV\zeta + 2\zeta^T V'^T QV\zeta, \qquad (3.2.16)$$

$$= 2\zeta'^T V^T QV\zeta + \zeta^T V'^T QV\zeta + \zeta^T V^T Q^T V'\zeta, \qquad (3.2.17)$$

$$\eta^T P\eta = \zeta^T V^T PV\zeta. \qquad (3.2.18)$$

Hence, $$2\omega = \zeta'^T R^*\zeta' + 2\zeta'^T Q^*\zeta + \zeta^T P^*\zeta, \qquad (3.2.19)$$

where

$$R^* = V^T R V, \tag{3.2.20}$$

$$Q^* = V^T Q V + V^T R V', \tag{3.2.21}$$

$$P^* = V^T P V + V'^T Q V + V^T Q^T V' + V'^T R V'. \tag{3.2.22}$$

Similarly, it can be shown that 2γ and Ψ^{μ} remain as quadratic and linear forms respectively. Thus it involves no loss of generality if it is assumed that R is of the form (3.2.11). By the use of such nonsingular linear transformations we may furthermore assume that the submatrix R_1 is nonsingular or does not occur.

3.3. THE GENERALIZED LEGENDRE CONDITION

Suppose the matrix R of the second variation of a given singular extremal E is of the form (3.2.11). Then we partition the matrices Q, P and η such that

$$Q \equiv \begin{pmatrix} Q_1 & Q_4 \\ Q_3 & Q_2 \end{pmatrix}, \tag{3.3.1}$$

$$P \equiv \begin{pmatrix} P_1 & P_3^T \\ P_3 & P_2 \end{pmatrix}, \tag{3.3.2}$$

$$\text{and} \quad \eta^T \equiv (\sigma^T \quad \kappa^T) , \tag{3.3.3}$$

where partition lines run between the mth and (m+1)th rows/columns.

<u>Theorem 1</u>. Along such a singular extremal E the following conditions are necessary:

(i) the $(n-m) \times (n-m)$ order matrix Q_2 must be identically symmetric,

(ii) if Q_2 is identically symmetric, then the matrix

$$R_4 \equiv \begin{pmatrix} R_1 & R_3^T \\ R_3 & R_2 \end{pmatrix}, \tag{3.3.4}$$

must be positive semidefinite where

$$R_3 \equiv Q_4^T - Q_3, \tag{3.3.5}$$

and
$$R_2 \equiv P_2 - Q_2'. \tag{3.3.6}$$

<u>Proof</u>: We shall first prove condition (ii). In the second variation the derivated functions $\kappa'_\rho(x)$, $\rho = m+1, m+2, \ldots, n$, occur only in the bilinear forms

$$2\kappa'^T Q_3 \sigma \text{ and } 2\kappa'^T Q_2 \kappa. \tag{3.3.7}$$

The assumption that Q_2 is symmetric leads to

$$\int_{x_1}^{x_2} 2\kappa'^T Q_2 \kappa dx = \kappa^T Q_2 \kappa \Big|_{x_1}^{x_2} - \int_{x_1}^{x_2} \kappa^T Q_2' \kappa dx \tag{3.3.8}$$

and in general,

$$\int_{x_1}^{x_2} 2\kappa'^T Q_3 \sigma dx = 2\kappa^T Q_3 \sigma \Big|_{x_1}^{x_2} - \int_{x_1}^{x_2} 2\kappa^T \frac{d}{dx}(Q_3\sigma) dx,$$

$$= 2\kappa^T Q_3 \sigma \Big|_{x_1}^{x_2} - \int_{x_1}^{x_2} 2[\kappa^T Q_3' \sigma + \kappa^T Q_3 \sigma'] dx. \tag{3.3.9}$$

Using equations (3.3.8) and (3.3.9) the $\kappa'_\rho(x)$ terms are completely eliminated from the second variation and we have

$$
\begin{aligned}
2\omega = {} & \sigma'^T R_1 \sigma' + 2\sigma'^T Q_1 \sigma + \sigma^T P_1 \sigma \\
& - 2\kappa^T Q_3' \sigma - 2\kappa^T Q_3 \sigma' - \kappa^T Q_2' \kappa \\
& + 2\sigma'^T Q_4 \kappa + 2\kappa^T P_3 \sigma + \kappa^T P_2 \kappa. \qquad (3.3.10)
\end{aligned}
$$

Thus the status of the $\kappa_\rho(x)$ terms can be raised to that of derivatives and this step is taken by replacing $\kappa_\rho(x)$ by $\kappa_\rho'^*(x)$ and after rearrangements

$$
\begin{aligned}
2\omega = {} & \sigma'^T R_1 \sigma' + 2\sigma'^T Q_1 \sigma + \sigma^T P_1 \sigma \\
& - 2\kappa'^{*T} Q_3 \sigma' + 2\kappa'^{*T} Q_4^T \sigma' \\
& - \kappa'^{*T} Q_2' \kappa'^* + \kappa'^{*T} P_2 \kappa'^* \\
& + 2\kappa'^{*T} P_3 \sigma - 2\kappa'^{*T} Q_3' \sigma. \qquad (3.3.11)
\end{aligned}
$$

This is of the same form as (3.2.4) but with

$$
R = \begin{pmatrix} R_1 & R_3^T \\ R_3 & R_2 \end{pmatrix}, \qquad (3.3.12)
$$

$$
Q \equiv \begin{pmatrix} Q_1 & 0 \\ Q_5 & 0 \end{pmatrix}, \qquad (3.3.13)
$$

$$
P \equiv \begin{pmatrix} P_1 & 0 \\ 0 & 0 \end{pmatrix}, \qquad (3.3.14)
$$

$$
\eta^T \equiv (\sigma^T \quad \kappa^{*T}), \qquad (3.3.15)
$$

where $R_3 \equiv Q_4^T - Q_3$,

$R_2 \equiv P_2 - Q_2'$,

and $Q_5 \equiv P_3 - Q_3'$, (3.3.16)

Following this it is observed that the terms $\kappa_\rho'^*(x_1)$ and $\kappa_\rho'^*(x_2)$ appear in

$$2\gamma[\xi_1,\sigma_r(x_1),\kappa_\rho'^*(x_1),\xi_2,\sigma_r(x_2),\kappa_\rho'^*(x_2)] \qquad (3.3.17)$$

and

$$\Psi^\mu[\xi_1,\sigma_r(x_1),\kappa_\rho'^*(x_1),\xi_2,\sigma_r(x_2),\kappa_\rho'^*(x_2)]. \qquad (3.3.18)$$

However, any restrictions placed upon these two sets of end values, $\kappa_\rho'^*(x_1)$ and $\kappa_\rho'^*(x_2)$, do not reduce the class of sets of admissible functions, for any such set can be made to satisfy such restrictions by infinitesimal adjustments over small neighbourhoods of the endpoints and these adjustments will only affect the integral in J_2 infinitesimally. As a consequence the quantities $\kappa_\rho'^*(x_1)$ and $\kappa_\rho'^*(x_2)$ can be treated as parameters playing roles similar to those of ξ_1 and ξ_2. For convenience ξ will be used to denote all these parameters.

Finally the application of the classical Legendre condition to the accessory extremal

$$\xi = 0, \quad \eta_i(x) \equiv 0 \qquad (3.3.19)$$

of the transformed accessory minimum problem, arrived at in the preceding paragraphs, leads to the condition that along the singular extremal E of the primary problem, the matrix R_4 of (3.3.12) must be positive semidefinite. Q.E.D.

Let us now prove condition (i). If Q_2 is not symmetric we introduce new partition lines running between the m^*th and (m^*+1)th rows/columns where $m < m^* \leqslant (n-1)$.

$$R \equiv \begin{pmatrix} R_1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \tag{3.3.20}$$

$$Q \equiv \begin{pmatrix} Q_1 & Q_4^* & Q_4^{**} \\ Q_3^* & Q_5 & Q_8 \\ Q_3^{**} & Q_7 & Q_6 \end{pmatrix}, \quad \text{say,} \tag{3.3.21}$$

$$P \equiv \begin{pmatrix} P_1 & P_3^{*T} & P_3^{**T} \\ P_3^* & P_5 & P_7^T \\ P_3^{**} & P_7 & P_6 \end{pmatrix}, \quad \text{say,} \tag{3.3.22}$$

Furthermore, m^* is chosen such that the $(n-m^*) \times (n-m^*)$ order matrix Q_6 is symmetric and is of the highest possible order. With m^* chosen in this manner, the non-symmetry of Q_2 implies that

$$Q_8^T - Q_7 \not\equiv 0. \tag{3.3.23}$$

This is easily seen on examining

$$Q_2 \equiv \begin{pmatrix} Q_5 & Q_8 \\ Q_7 & Q_6 \end{pmatrix}. \tag{3.3.24}$$

If $Q_8^T - Q_7 \equiv 0$, then the order of the symmetric matrix Q_6 can be increased by at least one. This is because the $(n-m^*+1) \times (n-m^*+1)$ matrix consisting of Q_6, the last column of Q_7, the last row of Q_8 and the last diagonal element of Q_5 would then be symmetric. In the extreme case Q_6 is an one $\times$ one order matrix which is then trivially symmetric. This accounts for $m^* \leqslant n-1$.

As Q_6 is symmetric, condition (ii) is applicable to the accessory minimum problem with matrices R, Q, P displayed in (3.3.20) to (3.3.22). Thus we are led to the condition that the matrix

$$R_4 \equiv \begin{pmatrix} R_1 & 0 & R_3^{*T} \\ 0 & 0 & R_3^{**T} \\ R_3^{*} & R_3^{**} & R_2^{*} \end{pmatrix}, \tag{3.3.25}$$

must be positive semidefinite where

$$\begin{aligned} (R_3^{*} \mid R_3^{**}) &\equiv (Q_4^{**T} \mid Q_8^{T}) - (Q_3^{**} \mid Q_7) \\ &= (Q_4^{**T} - Q_3^{**} \mid Q_8^{T} - Q_7), \end{aligned} \tag{3.3.26}$$

$$\Rightarrow \quad R_3^{*} = Q_4^{**T} - Q_3^{**} \tag{3.3.27}$$

$$\text{and} \quad R_3^{**} = Q_8^{T} - Q_7. \tag{3.3.28}$$

Since R_4 must be positive semidefinite along the singular extremal E, hence the $(n-m) \times (n-m)$ order submatrix

$$\begin{pmatrix} 0 & R_3^{**T} \\ R_3^{**} & R_2^{*} \end{pmatrix} \tag{3.3.29}$$

must be positive semidefinite. Hence by Lemma 1, the matrix

$$R_3^{**} \equiv 0. \tag{3.3.30}$$

From (3.3.23) and (3.3.28) this condition is not satisfied. Hence we conclude that Q_2 must be symmetric. Q.E.D.

<u>Corollary 1.1</u>. Along a singular extremal satisfying Theorem 1 and along which the matrices

$$R_2 \equiv 0 \tag{3.3.31}$$

and

$$R_3 \equiv 0 , \tag{3.3.32}$$

the matrix $Q_5 = P_3 - Q_3'$ must be equal to the zero matrix identically.

Proof: For such an extremal E the transformed second variation is in the same form as that displayed in (3.2.3) but with R, Q, P and η in the forms displayed in (3.2.11) and (3.3.13) to (3.3.15) respectively. As R is in the form (3.2.11) the same method as that used to prove Theorem 1 may be used to prove that

$$R_4 \equiv \begin{pmatrix} R_1 & -Q_5^T \\ -Q_5 & 0 \end{pmatrix} , \tag{3.3.33}$$

must be positive semidefinite. This is easily seen on comparing the matrices R, Q, P, η displayed in (3.3.11) and (3.3.13) to (3.3.15) with the corresponding matrices displayed in (3.2.11) and (3.3.1) to (3.3.3). Furthermore, note that in the transformed second variation the matrix corresponding to Q_2 of Theorem 1 is the zero matrix and is therefore trivially symmetric. Finally by Lemma 1 we are led to conclude that

$$Q_5 \equiv 0. \qquad \text{Q.E.D.} \tag{3.3.34}$$

Remarks: If a given singular extremal satisfies Theorem 1 and Corollary 1.1, the matrices of the R, P, η are then in the forms displayed in (3.3.11), (3.3.14) and (3.3.15). Furthermore

$$Q \equiv \begin{pmatrix} Q_1 & 0 \\ 0 & 0 \end{pmatrix} \tag{3.3.35}$$

and hence the variations $\kappa^*_\rho(x)$ do not occur in the integrand of the second variation. Therefore further necessary conditions must be deduced from an examination of 2γ and the end conditions and subsequently a mixture of the second variation and higher order variations must be studied. The latter would undoubtedly be very difficult. This was the difficulty that was expected by earlier researchers on singular extremals. It corresponds to the difficulties encountered in the semidefinite case in the problem of the minimization of a function of several variables in differential calculus, Hancock [29], Chaundy [30].

3.4. THE JACOBI CONDITION

Here we will propose ways by which necessary conditions of the Jacobi type may be imposed on a singular extremal E, which satisfies Theorem 1 above. Consider a fixed endpoints variational problem similar to that formulated in the preceding sections. Assuming that $y''_i(x)$ of E are continuous, the second variation is well defined and is of the form

$$J_2 = \int_{x_1}^{x_2} 2\omega(x,\eta,\eta')dx, \tag{3.4.1}$$

where $$2\omega = \eta'^T R\eta' + 2\eta'^T Q\eta + \eta^T P\eta, \tag{3.4.2}$$

$$\eta_i(x_1) = 0 = \eta_i(x_2), \tag{3.4.3}$$

and $$\eta_i(x) \in D' \text{ in } [x_1,x_2]. \tag{3.4.4}$$

Consider the case where R is of the form (3.2.11), and where $R_1(x)$ is a $m \times m$ order $(0 < m < n)$ nonsingular matrix which is positive definite. Next assume that the matrices Q, P, η are partitioned as shown in (3.3.1) to (3.3.3) respectively. Letting

$$\kappa(x) \equiv 0,$$

the second variation is reduced to

$$J_2 = \int_{x_1}^{x_2} \sigma'^T R_1 \sigma' + 2\sigma'^T Q_1 \sigma + \sigma^T P_1 \sigma dx. \qquad (3.4.5)$$

Definition: Let $U(x)$ be a $m \times m$ order matrix solution of the differential system

$$\frac{d}{dx}(R_1 U' + Q_1 U) = Q_1^T U' + P_1 U\,, \qquad (3.4.6)$$

where $U(x)$ is not identically singular in $[x_1, x_2]$. Then x_3 is said to be conjugate to x_1 relative to this differential system if the determinants

$$|U(x_1)| = 0 = |U(x_3)|. \qquad (3.4.7)$$

By means of the accessory minimum problem involving the second variation (3.4.5), we can impose the necessary condition, that there must not be a conjugate point of x_1, relative to the differential system (3.4.6) in the interior of $[x_1, x_2]$.

The geometrical interpretation is quite simple. In general we must first assume the existence of the singular extremal E, after which if $0 < m$ we may expect E to be imbedded in a family of extremals passing through the point on E where $x = x_1$. This family of extremals generates a hypersurface in contrast to the case when E is a regular extremal for which such a family of extremals fills the whole neighbourhood of E. When there exists a conjugate point in the above reduced sense it appears very plausible that there exists a one parameter family of extremals lying in this hypersurface and touching an envelope at the conjugate point. These observations remain an interesting research topic.

The complete Jacobi condition for this problem involves the transformed second variation with matrices R, Q, P, η displayed in (3.3.12) to (3.3.16) and where $\kappa_\rho^{\prime *}(x_1)$ and $\kappa_\rho^{\prime *}(x_2)$ are treated like parameters ξ_1, ξ_2. This is an accessory minimum problem with variable endpoints.

3.5. MINIMIZATION OF A SINGULAR QUADRATIC FUNCTIONAL

Consider the following variational problem with fixed end points:

$$\min. \int_{x_1}^{x_2} y'^T R y' + 2y'^T Q y + y^T P y dx, \tag{3.5.1}$$

where

$$R(x) \equiv \begin{pmatrix} R_1 & 0 \\ 0 & 0 \end{pmatrix}, \tag{3.5.2}$$

$$Q(x) \equiv \begin{pmatrix} Q_1 & Q_4 \\ Q_3 & Q_2 \end{pmatrix}, \tag{3.5.3}$$

$$P(x) \equiv \begin{pmatrix} P_1 & P_3^T \\ P_3 & P_2 \end{pmatrix}, \qquad (3.5.4)$$

$$y^T \equiv (u^T \quad v^T), \qquad (3.5.5)$$

$$y_i(x_1) = a_i, \quad y_i(x_2) = b_i, \qquad (3.5.6)$$

$$i = 1,2,\ldots,n: \quad j = 1,2,\ldots,m \ (0 < m < n): \quad \rho = m+1, m+2, \ldots, n; \qquad (3.5.7)$$

$$Q_2 \equiv Q_2^T \text{ (i.e. } Q_2 \text{ is symmetric)}, \qquad (3.5.8)$$

and
$$R_4 \equiv \begin{pmatrix} R_1 & Q_4 - Q_3^T \\ Q_4^T - Q_3 & P_2 - Q_2' \end{pmatrix}, \qquad (3.5.9)$$

is positive definite in $[x_1, x_2]$.

In matrix notation, the Euler equation is

$$(d/dx)(Ry' + Qy) = Q^T y' + Py, \qquad (3.5.10)$$

$$\therefore \; Ry'' + R'y + Qy' + Q'y = Q^T y' + Py,$$
$$Ry'' + (Q - Q^T)y' + (R' + Q' - P)y = 0. \qquad (3.5.11)$$

Using (3.5.2) to (3.5.8) we have

$$\begin{pmatrix} R_1 & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} u'' \\ v'' \end{pmatrix} + \begin{pmatrix} Q_1 - Q_1^T & Q_4 - Q_3^T \\ Q_3 - Q_4^T & 0 \end{pmatrix}\begin{pmatrix} u' \\ v' \end{pmatrix}$$
$$+ \begin{pmatrix} R_1' + Q_1' - P_1 & Q_4' - P_3^T \\ Q_3' - P_3 & Q_2' - P \end{pmatrix}\begin{pmatrix} u \\ v \end{pmatrix} = 0. \qquad (3.5.12)$$

$$\therefore \; R_1 u'' + (Q_1 - Q_1^T)u' + (R_1' + Q_1' - P_1)u$$
$$+ (Q_4 - Q_3^T)v' + (Q_4' - P_3^T)v = 0,$$
$$(Q_3 - Q_4^T)u' + (Q_3' - P_3)u + (Q_2' - P_2)v = 0.$$

Introducing the notations,

$$R_3(x) \equiv Q_4^T - Q_3, \quad R_2 \equiv (P_2 - Q_2'), \quad Q_5 \equiv (P_3 - Q_3'), \qquad (3.5.13)$$

we have

$$R_1u'' + (Q_1 - Q_1^T)u' + (R_1' + Q_1' - P_1)u + R_3^Tv' + (Q_4' - P_3^T)v = 0, \qquad (3.5.14)$$

and

$$R_3u' + Q_5u + R_2v = 0. \qquad (3.5.15)$$

The assumption that R_4 is positive definite implies that R_2 is also positive definite and consequently is nonsingular. Hence (3.5.15) yields

$$v + R_2^{-1}R_3u' + R_2^{-1}Q_5u = 0. \qquad (3.5.16)$$

Differentiating, we have

$$v' + R_2^{-1}R_3u'' + [(d/dx)(R_2^{-1}R_3) + R_2^{-1}Q_5]u' + (d/dx)(R_2^{-1}Q_5)u = 0. \qquad (3.5.17)$$

Substituting (3.5.16) and (3.5.17) into (3.5.14) and rearranging, we get

$$(R_1 - R_3^TR_2^{-1}R_3)u'' + [Q_1 - Q_1^T - R_3^T(d/dx)(R_2^{-1}R_3) - R_3^TR_2^{-1}Q_5 - (Q_4' - P_3^T)R_2^{-1}R_3]u' + [R_1' + Q_1' - P_1 - R_3^T(d/dx)(R_2^{-1}Q_5) - (Q_4' - P_3^T)R_2^{-1}Q_5]u = 0. \qquad (3.5.18)$$

The positive definiteness of R_4 of (3.5.9) implies that the $m \times m$ matrix $(R_1 - R_3^TR_2^{-1}R_3)$ is also positive definite, Gantmacher [31, p.46], and consequently is nonsingular. Hence

(3.5.18) is a regular differential system whose general solution has 2m arbitrary constants. The procedure for the derivation of the singular extremals is now clear. Firstly solve (3.5.18) for the m-order vector u subjected to the end conditions $y_j(x_1)$ and $y_j(x_2)$. Then v is obtained from (3.5.16). In general, this vector v will not satisfy the end conditions imposed on $y_\rho(x_1)$ and $y_\rho(x_2)$.

A different approach involves differentiating (3.5.15) which gives after rearrangements,

$$R_3u'' + R_2v' + R_2v + (R_3' + Q_5)u' + Q_5u = 0. \qquad (3.5.19)$$

This equation and equation (3.5.14) form a regular differential system and the matrix coefficient of $\begin{pmatrix} u'' \\ v' \end{pmatrix}$ is

$$\begin{pmatrix} R_1 & R_3^T \\ R_3 & R_2 \end{pmatrix} \qquad (3.5.20)$$

which is R_4. By hypothesis R_4 is positive definite and consequently is nonsingular. Hence the general solution of this differential system contains $(2m + n - m)$ arbitrary constants, $(n - m)$ of which must subsequently be determined from (3.5.16).

3.6. EXTREMIZATION OF THE GENERAL LINEAR INTEGRAL.

Consider the variational problem with fixed end points:

$$\min.\int_{x_1}^{x_2} L_0(x,y) + y_i'L_i(x,y)dx\,, \tag{3.6.1}$$

where
$$y_i(x_1) = a_i, \quad y_i(x_2) = b_i, \tag{3.6.2}$$

and $L_0(x,y)$, $L_i(x,y)$ are known functions of (x,y). We will employ the indices

$$i,j,r = 1,2,\dots,n. \tag{3.6.3}$$

The Euler equations are

$$\frac{d}{dx}L_i = y_j'\frac{\partial L_j}{\partial y_i} + \frac{\partial L_0}{\partial y_i}\,, \tag{3.6.4}$$

$$\therefore \quad y_j'\left(\frac{\partial L_j}{\partial y_i} - \frac{\partial L_i}{\partial y_j}\right) + \frac{\partial L_0}{\partial y_i} - \frac{\partial L_i}{\partial x} = 0. \tag{3.6.5}$$

Assuming that there exists a smooth arc E satisfying (3.6.5), the Weierstrass condition requires that

$$y_i'^*L_i - y_i'L_i - (y_i'^* - y_i')L_i \geq 0 \tag{3.6.6}$$

and this is trivially satisfied along E. It can also be shown that the classical Legendre condition is also trivially satisfied.

Under the further assumption that along E the functions $y_i''(x)$ exist and are continuous, the second variation is well defined and is of the form

$$J_2 \equiv \int_{x_1}^{x_2} \eta'^T R\eta' + 2\eta'^T Q\eta + \eta^T P\eta dx, \tag{3.6.7}$$

where
$$R(x) \equiv 0, \tag{3.6.8}$$

$$Q(x) \equiv (\partial L_i/\partial y_j), \tag{3.6.9}$$

$$P(x) \equiv (\partial^2 L_0/\partial y_i \partial y_j + y_r' \partial^2 L_r/\partial y_i \partial y_j), \tag{3.6.10}$$

$$\eta_i(x_1) = 0 = \eta_i(x_2). \tag{3.6.11}$$

Hence Theorem 1 is applicable with $R_1 = 0$, $Q_2 = Q$ and $P_2 = P$. Therefore Q of (3.6.9) must be symmetric i.e.

$$\frac{\partial L_i}{\partial y_j} - \frac{\partial L_j}{\partial y_i} = 0\,. \tag{3.6.12}$$

Substituting these equations into (3.6.5), we get

$$\frac{\partial L_0}{\partial y_i} - \frac{\partial L_i}{\partial x} = 0. \tag{3.6.13}$$

Thus we have $n(n-1)/2+n=n(n+1)/2$ finite equations in (x,y) and in general there is no solution. Equations (3.6.12) and and (3.6.13) are also the well known necessary and sufficient conditions that the integral (3.6.1) be independent of the path of integration in a simply connected region of the (x,y) space.

From Theorem 1, along an extremal E satisfying (3.6.12) and (3.6.13) the matrix

$$R_4 = R_2 \equiv \left(\frac{\partial^2 L_0}{\partial y_i \partial y_j} + y_r' \frac{\partial^2 L_r}{\partial y_i \partial y_j} - \frac{d}{dx}\frac{\partial L_i}{\partial y_j}\right), \tag{3.6.14}$$

must be positive semidefinite. It can also be written thus:

$$R_2 \equiv \left(\frac{\partial^2 L_0}{\partial y_i \partial y_j} - \frac{\partial^2 L_i}{\partial x \partial y_j} + y_r' \frac{\partial^2 L_r}{\partial y_i \partial y_j} - y_r' \frac{\partial^2 L_i}{\partial y_j \partial y_r}\right). \tag{3.6.15}$$

A procedure for the derivation of the singular extremals is to examine firstly the $n(n-1)/2$ equations (3.6.12) for functions $y_i(x)$. For $n > 3$ there are more equations than variables and hence, in general, solutions $y_i(x)$ do not exist. Assuming that solutions $y_i(x)$ to these equations exist we next examine the n equations of (3.6.13). From the implicit functions theorem, the condition that the functions $y_i(x)$ can be uniquely determined from (3.6.13) is that the $n \times n$ matrix

$$\left(\frac{\partial^2 L_0}{\partial y_i \partial y_j} - \frac{\partial^2 L_i}{\partial x \partial y_j} \right) \qquad (3.6.16)$$

must be nonsingular.

We will now demonstrate the relationship between this matrix and the matrix R_2 of (3.6.15). Multiplying (3.6.12) by n undetermined functions $\lambda_i(x)$ and summing, we get n equations

$$\lambda_r \frac{\partial L_r}{\partial y_j} - \lambda_r \frac{\partial L_j}{\partial y_r} = 0. \qquad (3.6.17)$$

Later on we will equate the $\lambda_i(x)$ to $y_i'(x)$. Adding these n equations to (3.6.13) we get

$$\frac{\partial L_0}{\partial y_j} - \frac{\partial L_j}{\partial x} + \lambda_r \frac{\partial L_r}{\partial y_j} - \lambda_r \frac{\partial L_j}{\partial y_r} = 0. \qquad (3.6.18)$$

From the implicit functions theorem, the condition that these n equations can be solved for $y_i(x,\lambda)$ is that the matrix

$$R_2^* \equiv \left(\frac{\partial^2 L_0}{\partial y_i \partial y_j} - \frac{\partial^2 L_j}{\partial y_i \partial x} + \lambda_r \frac{\partial^2 L_r}{\partial y_i \partial y_j} - \lambda_r \frac{\partial^2 L_j}{\partial y_i \partial y_r} \right), \qquad (3.6.19)$$

must be nonsingular. However because of (3.6.12), these n equations can be made to be devoid of the functions $\lambda_i(x)$ and consequently it must be possible to simplify the solutions $y_i(x,\lambda)$ of (3.6.18) so that they are also devoid of the functions $\lambda_i(x)$. If we let

$$\lambda_i(x) = y_i'(x), \tag{3.6.20}$$

the matrix R_2^* of (3.6.19) becomes R_2 of (3.6.15). Hence if R_2 is positive definite and consequently is nonsingular and if furthermore the equations (3.6.12) and (3.6.13) are consistent then we are assured of unique solutions of $y_i(x)$ from the n equations (3.6.13).

3.7. HAMILTON'S PRINCIPLE.

Consider the Hamilton's principle in analytical mechanics. For a conservative system it requires that the system moves so as to minimize (over short interval of time) the integral

$$I = \int_{t_1}^{t_2} L(q,\dot{q})dt, \tag{3.7.1}$$

where $L(q,\dot{q})$ is the Lagrangian of the system. Consequently the equations of motion are the well known Lagrange equations

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_i} = \frac{\partial L}{\partial q_i}, \quad (i,j = 1,2,\ldots,n). \tag{3.7.2}$$

Canonical variables are introduced by defining

$$p_i = \frac{\partial L}{\partial \dot{q}_i} \tag{3.7.3}$$

and if the problem is nonsingular the matrix $(\partial^2 L/\partial \dot{q}_i \partial \dot{q}_j)$ is nonsingular and (3.7.3) can be solved for $\dot{q}_i = \psi_i(q,p)$, say. The Hamiltonian function is then well defined and it is given as

$$H(q,p) = p_i\dot{q}_i - L(q,\dot{q}), \tag{3.7.4}$$

$$= p_i\psi_i - L(q,\psi). \tag{3.7.5}$$

The equations of motion may then be written thus:

$$\dot{q}_i = \frac{\partial H}{\partial p_i}\ , \quad \dot{p}_i = -\frac{\partial H}{\partial q_i}\ . \tag{3.7.6}$$

These equations may also be obtained by requiring that the integral

$$I = \int_{t_1}^{t_2} p_i\dot{q}_i - H(q,p)dt, \tag{3.7.7}$$

$$= \int_{t_1}^{t_2} F(q,\dot{q},p)dt, \tag{3.7.8}$$

is stationary in the (2n+1)-dimensional space (t,q,p). In fact (3.7.6) are the Euler equations of (3.7.7). However this latter variational problem is singular for

$$\begin{vmatrix} \dfrac{\partial^2 F}{\partial\dot{q}_i\partial\dot{q}_j} & \dfrac{\partial^2 F}{\partial\dot{q}_i\partial\dot{p}_s} \\ \dfrac{\partial^2 F}{\partial\dot{p}_r\partial\dot{q}_j} & \dfrac{\partial^2 F}{\partial\dot{p}_r\partial\dot{p}_s} \end{vmatrix} \equiv 0, \quad (r,s = 1,2,\ldots,n). \tag{3.7.9}$$

Therefore it is not surprising that the general solution of the differential system (3.7.6) has only 2n instead of 4n arbitrary constants.

In the notation of Theorem 1, the matrix R_1 is nonexisting and

$$Q_2 \equiv \begin{pmatrix} \dfrac{\partial^2 F}{\partial \dot{q}_i \partial q_j} & \dfrac{\partial^2 F}{\partial \dot{q}_i \partial p_s} \\ \dfrac{\partial^2 F}{\partial \dot{p}_r \partial q_j} & \dfrac{\partial^2 F}{\partial \dot{p}_r \partial p_s} \end{pmatrix} = \begin{pmatrix} 0 & \delta_{is} \\ 0 & 0 \end{pmatrix}, \qquad (3.7.10)$$

where δ_{is} are the Kronecker deltas. Now condition (i) of Theorem 1 requires that along any of these singular extremals (3.7.6), the matrix Q_2 is symmetric and this is obviously not satisfied. Thus we conclude that the latter variational principle, sometimes known as the modified Hamilton's principle, is non-minimal, which is in agreement with the conclusions of Hestenes [26, sec. 4.8] and Block [32], who employed different methods.

CHAPTER 4. OPTIMAL CONTROL PROBLEMS

4.1. A REVIEW OF THE CLASSICAL NECESSARY CONDITIONS

In this chapter the following set of indices will be employed unless otherwise stated:

$$\begin{aligned} &i,j = 1,2,\ldots,n; \quad r,s = 1,2,\ldots,m; \\ &\alpha,\beta = 1,2,\ldots,m^* < m; \quad \rho,\nu = m^*+1, m^*+2,\ldots m; \\ &\mu = 1,2,\ldots,N \leq n+1. \end{aligned} \tag{4.1.1}$$

In current optimal control problems the following Bolza problem is of fundamental importance: Find the control functions $u_r(t)$ which minimizes the performance index

$$J \equiv g[x(t_1),t_1] + \int_{t_0}^{t_1} f(x,u,t)dt, \tag{4.1.2}$$

where the functions $x_i(t)$, $u_r(t)$ are subjected to the conditions

$$\dot{x}_i = f_i(x,u,t), \tag{4.1.3}$$

$$x_i(t_0) = x_{io}\,(\text{constants}), \tag{4.1.4}$$

$$\psi^{\mu}[x(t_1);t_1] = 0, \tag{4.1.5}$$

$$a_r \leq u_r \leq b_r \quad (a_r, b_r \text{ are constants}) . \tag{4.1.6}$$

However, unless otherwise stated we will assume that the control vector (u_r) belongs to an open region U i.e. $a_r < u_r < b_r$. After deriving the main results we will outline the necessary modifications to include the cases when some of the u_r's are equal to the corresponding a_r's and b_r's along the reference extremal.

This problem becomes an equivalent Bolza problem as formulated in chapter 1 above if we introduce m auxiliary variables $z_r(t)$ and transform the problem by eliminating the u_r's using the equations

$$u_r = \dot{z}_r, \quad z_r(t_o) = 0 \tag{4.1.7}$$

and assuming that the functions $z_r(t_1)$ are unconstrained. The initial values of $z_r(t)$ have been put equal to zero for definiteness and this step is of no consequence.

From experience it has been found that the necessary conditions can be stated more simply in terms of a Hamiltonian function H (pseudo-Hamiltonian function, Lagrange expression). Assuming that the problem is <u>normal</u> the Hamiltonian function may be defined by

$$H(\lambda,x,u,t) \equiv \lambda_i f_i(x,u,t) + f(x,u,t), \tag{4.1.8}$$

or

$$H(\lambda,x,u,t) \equiv \lambda_i f_i(x,u,t) - f(x,u,t). \tag{4.1.9}$$

To translate the necessary conditions of the standard Bolza problem into conditions involving the function $H(\lambda,x,u,t)$ the function $F(\lambda,x,u,t)$ of the Multiplier Rule is defined thus:

$$F \equiv f + \lambda_i(f_i - \dot{x}_i) , \tag{4.1.10}$$

$$= H - \lambda_i \dot{x}_i , \tag{4.1.11}$$

or $$F \equiv f + \lambda_i(\dot{x}_i - f_i) \;, \tag{4.1.12}$$

$$= \lambda_i \dot{x}_i - H \;, \tag{4.1.13}$$

depending on whether H is defined by (4.1.8) or (4.1.9). In this dissertation we will employ the Hamiltonian function H defined by (4.1.9). This definition conforms with the manner in which the Hamiltonian function has been traditionally defined. For Mayer problems the difference between these alternative Hamiltonian functions becomes apparent only on examination of the transversality conditions or the Weierstrass condition or the Clebsch condition.

It is readily shown that the necessary conditions for a smooth arc to be minimizing, are:

(i) <u>The Euler-Lagrange equations</u>

$$\dot{\lambda}_i = -\frac{\partial H}{\partial x_i} \tag{4.1.14}$$

$$\frac{\partial H}{\partial u_r} = 0 \;. \tag{4.1.15}$$

(ii) <u>The Transversality condition</u>

$$dg + e^{\mu} d\psi^{\mu} - H_1 dt_1 + \lambda_{i1} dx_{i1} = 0 \tag{4.1.16}$$

for all arbitrary values of dx_{i1} and dt_1. The constants e^{μ} are Lagrange multipliers.

(iii) <u>The Weierstrass condition</u>

$$H(\lambda,x,u,t) \geqslant H(\lambda,x,u^*,t) \qquad (4.1.17)$$

for all admissible u_r^* i.e. $(u_r^*) \in U$ and for every element (λ,x,u,t) of the reference extremal.

(iv) <u>The classical Clebsch condition</u>

$$-\frac{\partial^2 H}{\partial u_r \partial u_s} \pi_r \pi_s \geqslant 0 \qquad (4.1.18)$$

for all arbitrary values of π_r.

Hereafter we will take the point of view that the Clebsch condition is a direct consequence of the Weierstrass condition, Bliss [2,p.224]. The Clebsch condition is then applicable to singular extremals. Finally we observe that corner (junction) conditions do not occur because of the assumption that the reference arc is smooth.

4.2. THE SINGULAR SECOND VARIATION

Let the (n+m) functions $y_i(t)$, $v_r(t)$ be the variations of $x_i(t)$ and $z_r(t)$ along the minimizing arc E and let ξ_1 be the variation of t_1. Assuming that the second derivatives $\ddot{x}_i(t)$ $\ddot{z}_r(t)$ exist and are continuous along E, the second variation of the functional J is well defined and is expressible in the form, Bliss [2,p.227],

$$J_2 \equiv 2\gamma[\xi_1, y(t_1)] + \int_{t_0}^{t_1} 2\omega(t, y, \dot{v})dt, \qquad (4.2.1)$$

in which $2\gamma[\xi_1, y(t_1)]$ is a homogeneous quadratic form in its arguments and

$$2\omega \equiv - \frac{\partial^2 H}{\partial u_r \partial u_s} \dot{v}_r \dot{v}_s - 2 \frac{\partial^2 H}{\partial u_r \partial x_i} \dot{v}_r y_i - \frac{\partial^2 H}{\partial x_i \partial x_j} y_i y_j . \qquad (4.2.2)$$

The equations of variation are

$$\dot{y}_i = \frac{\partial f_i}{\partial x_j} y_j + \frac{\partial f_i}{\partial u_r} \dot{v}_r , \qquad (4.2.3)$$

$$y_i(t_o) = 0, \quad v_r(t_o) = 0 , \qquad (4.2.4)$$

$$\Psi^{\mu} [\xi_1, y(t_1)] = 0, \qquad (4.2.5)$$

where $v_r(t_1)$ are unconstrained and $\Psi^{\mu}[\xi_1, y(t_1)]$ stand for N sets of linear homogeneous forms in its arguments.

In matrix notation 2ω and the constraints have the form

$$2\omega \equiv \dot{v}^T R \dot{v} + 2\dot{v}^T Q y + y^T P y , \qquad (4.2.6)$$

$$\dot{y} = Ay + B\dot{v}, \qquad (4.2.7)$$

where

$$R \equiv \left(- \frac{\partial^2 H}{\partial u_r \partial u_s}\right), \qquad (4.2.8)$$

$$Q \equiv \left(- \frac{\partial^2 H}{\partial u_r \partial x_i}\right), \qquad (4.2.9)$$

$$P \equiv \left(- \frac{\partial^2 H}{\partial x_i \partial x_j}\right), \qquad (4.2.10)$$

$$A \equiv \left(\frac{\partial f_i}{\partial x_j}\right) , \qquad (4.2.11)$$

$$B \equiv \left(\frac{\partial f_i}{\partial u_r}\right) . \tag{4.2.12}$$

From (1.3.9) an extremal E is said to be singular if the $(2n+m) \times (2n+m)$ order determinant

$$\Delta \equiv \begin{vmatrix} 0 & 0 & I \\ 0 & R & B^T \\ I & B & 0 \end{vmatrix} , \tag{4.2.13}$$

vanishes identically along E. Expanding this determinant by the first n columns followed by the first n rows we get

$$\begin{aligned} \Delta &= (-1)^{(n+m+2)n} \times (-1)^{(m+2)n} |R| \\ &= (-1)^{n^2} |R| \ . \end{aligned}$$

Hence for this class of Bolza problems we can say that an extremal is singular if the determinant

$$\left| - \frac{\partial^2 H}{\partial u_r \partial u_s} \right| \equiv 0, \tag{4.2.14}$$

along the extremal.

We will now show that, without loss of generality, we may assume that along a singular extremal E, the matrix R has the form

$$R \equiv \begin{pmatrix} R_1 & 0 \\ 0 & 0 \end{pmatrix}, \tag{4.2.15}$$

where R_1 is a $m^* \times m^*$ order nonsingular submatrix and where

$$0 \leqslant m^* < m \ . \tag{4.2.16}$$

When $m^* = 0$, R_1 does not occur.

Firstly an important class of singular Bolza problems consists of problems in which one or more control variables appear linearly in (4.1.2) and (4.1.3) and in which the classical Clebsch condition is satisfied. Thus if the control variable u_m, say, appears linearly we have

$$\frac{\partial^2 H}{\partial u_m^2} \equiv 0 \tag{4.2.17}$$

and the matrix R is of the form

$$R \equiv \begin{pmatrix} -\dfrac{\partial^2 H}{\partial u_r \partial u_s} & -\dfrac{\partial^2 H}{\partial u_r \partial u_m} \\ -\dfrac{\partial^2 H}{\partial u_s \partial u_m} & 0 \end{pmatrix} , \tag{4.2.18}$$

for $r,s = 1,2,\ldots,m-1$. The classical Clebsch condition requires the matrix R to be positive semidefinite and hence by Lemma 1 the $1 \times (m-1)$ order matrix

$$\left(-\frac{\partial^2 H}{\partial u_s \partial u_m} \right) \equiv 0 \ . \tag{4.2.19}$$

Therefore R is of the form (4.2.15).

In a similiar manner when $(m-m^*)$ control variables appear linearly in both (4.1.2) and (4.1.3) and when the classical Clebsch condition is satisfied the $m \times m$ matrix R is expressible in the form displayed in (4.2.15) along the reference extremal, where R_1 is a $m^* \times m^*$ order matrix. Most of the singular Bolza problems of optimal control theory belong to this class.

For other singular extremals it may be necessary to subject the accessory minimum problem to a certain linear transformation before we can assume that the matrix R is of the form (4.2.15). From matrix theory we know that, as R is symmetric and singular, then there exists a nonsingular square matrix V such that

$$V^T R V = \begin{pmatrix} R_1 & 0 \\ 0 & 0 \end{pmatrix}, \qquad (4.2.20)$$

where R_1 is nonsingular or nonexisting. Employing such a matrix V we subject the vector $\dot{v}$ to the regular transformation

$$\dot{v} = V\dot{w}, \text{ say.} \qquad (4.2.21)$$

Under such a transformation

$$2\omega(t,y,\dot{w}) = \dot{w}^T V^T R V \dot{w} + 2\dot{w}^T V^T Q y + y^T P y \qquad (4.2.22)$$

and

$$\dot{y} = Ay + BV\dot{w} \qquad (4.2.23)$$

and these are of the same form as (4.2.6) and (4.2.7). Therefore, without loss of generality, we can assume that along a smooth singular extremal E the matrix R is of the form (4.2.15).

4.3. THE GENERALISED CLEBSCH CONDITION

Suppose that along a singular extremal E, the matrix R of the second variation is of the form

$$R \equiv \begin{pmatrix} R_1 & 0 \\ 0 & 0 \end{pmatrix} , \tag{4.3.1}$$

where R_1 is a $m^* \times m^*$ nonsingular matrix $(0 \leqslant m^* < m)$. Then we partition the matrices Q,B and v in a similar manner, that is,

$$Q \equiv \begin{pmatrix} Q_1 \\ Q_2 \end{pmatrix} , \tag{4.3.2}$$

$$B \equiv (B_1 \quad B_2) , \tag{4.3.3}$$

$$v \equiv \begin{pmatrix} \sigma \\ \kappa \end{pmatrix} , \tag{4.3.4}$$

where Q_1 is a $m^* \times n$ matrix, Q_2 is a $(m-m^*) \times n$ matrix, B_1 is a $n \times m^*$ matrix, B_2 is a $n \times (m-m^*)$ matrix, σ is a $m^* \times 1$ matrix and κ is a $(m-m^*) \times 1$ matrix.

Theorem 2

If along a singular minimizing arc E the matrices R,Q,B are partitioned in the manner displayed in (4.3.1) to (4.3.3) then along arc E the following conditions are necessary:

(i) The $(m-m^*) \times (m-m^*)$ order matrix $Q_2 B_2$ must be identically symmetric.

(ii) If $Q_2 B_2$ is identically symmetric then the matrix

$$\begin{pmatrix} R_1 & R_2^T \\ R_2 & R_3 \end{pmatrix} \equiv R_4 \text{ say,} \qquad (4.3.5)$$

must be positive semidefinite where

$$R_2 \equiv B_2{}^T Q_1{}^T - Q_2 B_1, \qquad (4.3.6)$$

$$R_3 \equiv B_2{}^T P B_2 - \frac{d}{dt}(Q_2 B_2) - Q_2 B_3 - B_3{}^T Q_2{}^T, \qquad (4.3.7)$$

and

$$B_3 \equiv AB_2 - \dot{B}_2 \;. \qquad (4.3.8)$$

Proof

We shall first prove condition (ii); that is we assume that $Q_2 B_2$ is symmetric. Before the method used to prove theorem 1 above can be employed, we must first eliminate the derivated vector $\dot{\kappa}$ from the equation of variation (4.2.7), which can also be written thus,

$$\dot{y} = Ay + B_1 \dot{\sigma} + B_2 \dot{\kappa}. \qquad (4.3.9)$$

A transformation that does this is

$$y = \tau + B_2 \kappa \;, \qquad (4.3.10)$$

where τ is a $n \times 1$ matrix. The inverse of this transformation is obviously,

$$\tau = y - B_2 \kappa \;. \qquad (4.3.11)$$

Therefore the transformation (4.3.10) is a one-to-one mapping of the (y,v)-space into the (τ,v)-space. Another motivation for the transformation (4.3.10) is that it leaves 2ω of (4.2.6) in a form in which $\dot{\kappa}$ may possibly be eliminated from the second variation by means of an integration by parts.

Differentiating (4.3.10) we get

$$\dot{y} = \dot{\tau} + B_2\dot{\kappa} + \dot{B}_2\kappa \ . \tag{4.3.12}$$

Using (4.3.10) and (4.3.12), we have

$$2\omega = \dot{\sigma}^T R_1\dot{\sigma} + 2\dot{\sigma}^T Q_1 y + y^T P y + 2\dot{\kappa}^T Q_2 y, \tag{4.3.13}$$

$$\begin{aligned} &= \dot{\sigma}^T R_1\dot{\sigma} + 2\dot{\sigma}^T Q_1\tau + 2\dot{\sigma}^T Q_1 B_2\kappa + \kappa^T B_2^T P B_2\kappa + \kappa^T B_2^T P\tau \\ &\qquad + \tau^T P B_2\kappa + \tau P\tau + 2\dot{\kappa}^T Q_2\tau + 2\dot{\kappa}^T Q_2 B_2\kappa, \end{aligned} \tag{4.3.14}$$

and $$\dot{\tau} + B_2\dot{\kappa} + \dot{B}_2\kappa = A\tau + AB_2\kappa + B_1\dot{\sigma} + B_2\dot{\kappa}$$

$$\therefore \quad \dot{\tau} = A\tau + B_1\dot{\sigma} + (AB_2 - \dot{B}_2)\kappa, \tag{4.3.15}$$

$$= A\tau + B_1\dot{\sigma} + B_3\kappa \ . \tag{4.3.16}$$

Hence we have successfully eliminated $\dot{\kappa}$ from the equation of variation.

On examination of (4.3.14), it is seen that $\dot{\kappa}$ occurs in 2ω only in the bilinear forms

$$2\dot{\kappa}^T Q_2\tau \quad \text{and} \quad 2\dot{\kappa}^T Q_2 B_2\kappa. \tag{4.3.17}$$

The assumption that Q_2B_2 is symmetric leads to

$$\int_{t_0}^{t_1} 2\dot{\kappa}^T Q_2 B_2\kappa dt = \kappa^T Q_2 B_2\kappa \Big|_{t_0}^{t_1} - \int_{t_0}^{t_1} \kappa \frac{d}{dt}(Q_2B_2)\kappa dt \tag{4.3.18}$$

and in general

$$\int_{t_0}^{t_1} 2\dot{\kappa}^T Q_2 \tau dt = 2\kappa^T Q_2 \tau \Big|_{t_0}^{t_1} - \int_{t_0}^{t_1} 2\kappa^T \frac{d}{dt}(Q_2\tau) dt \ . \qquad (4.3.19)$$

Employing the equation of variation (4.3.16), we have

$$\begin{aligned} -2\kappa^T(d/dt)(Q_2\tau) &= -\ 2\kappa^T\dot{Q}_2\tau - 2\kappa^T Q_2\dot{\tau} \\ &= -\ 2\kappa^T\dot{Q}_2\tau - 2\kappa^T Q_2 B_1\dot{\sigma} - 2\kappa^T Q_2 A\tau - 2\kappa^T Q_2 B_3\kappa. \qquad (4.3.20) \end{aligned}$$

Substituting (4.3.18) to (4.3.20) into (4.3.14) we get

$$\begin{aligned} 2\omega = {} & \dot{\sigma}^T R_1\dot{\sigma} + 2\dot{\sigma}^T Q_1 B_2\kappa - 2\kappa^T Q_2 B_1\dot{\sigma} + \kappa^T\Big[B_2^T P B_2 - (d/dt)(Q_2 B_2) - 2Q_2 B_3\Big]\kappa \\ & + 2\dot{\sigma}^T Q_1\tau + \kappa^T B_2 P\tau + \tau^T P B_2\kappa - 2\kappa^T\dot{Q}_2\tau - 2\kappa^T Q_2 A\tau + \tau^T P\tau. \qquad (4.3.21) \end{aligned}$$

Employing the notations (4.3.6) and (4.3.7) and

$$Q_3 \equiv B_2^T P - \dot{Q}_2 - Q_2 A \qquad (4.3.22)$$

and after simplifications we have

$$2\omega = \dot{\sigma}^T R_1\dot{\sigma} + 2\kappa^T R_2\dot{\sigma} + \kappa^T R_3\kappa + 2\dot{\sigma}^T Q_1\tau + 2\kappa^T Q_3\tau + \tau^T P\tau. \qquad (4.3.23)$$

Hence the vector $\dot{\kappa}$ is completely eliminated from the second variation and the equation of variation. Therefore the status of the vector κ, in the transformed accessory minimum problem, can be raised to that of a derivated vector and this step is taken by replacing κ with $\dot{\kappa}^*$, say.

Following this, it is observed that the vector $\dot{\kappa}^*$ appears in

$$2\gamma[\xi,\tau(t_1),\dot{\kappa}^*(t_1)], \quad \Psi^{\mu}[\xi,\tau(t_1),\dot{\kappa}^*(t_1)]. \qquad (4.3.24)$$

However, any restrictions placed upon the vector $\dot{\kappa}^*(t_1)$ do not reduce the class of sets of admissible vector functions $\tau(t)$, $\sigma(t)$, $\kappa^*(t)$, satisfying the matrix equation of variation, for any such set can be made to satisfy such restrictions by infinitesimal adjustments over small neighbourhoods of the endpoints and these adjustments will only affect the integral in J_2 infinitesimally. As a consequence, the quantities $\dot{\kappa}^*_\rho(t_1)$ can be treated as parameters playing roles similar to that of ξ_1. Hereafter we will let ξ denote the set ξ_1 and $\dot{\kappa}^*_\rho(t_1)$.

Employing the transformation

$$\kappa = \dot{\kappa}^* \tag{4.3.25}$$

the equation of variation becomes

$$\dot{\tau} = A\tau + B_1\dot{\sigma} + B_3\dot{\kappa}^* \tag{4.3.26}$$

and

$$2\omega = \dot{\sigma}^T R_1 \dot{\sigma} + 2\dot{\kappa}^{*T} R_2 \dot{\sigma} + \dot{\kappa}^{*T} R_3 \dot{\kappa}^* + 2\dot{\sigma}^T Q_1 \tau + 2\dot{\kappa}^{*T} Q_3 \tau + \tau^T P \tau \; . \tag{4.3.27}$$

In this transformed accessory minimum problem the application of the classical Clebsch condition to the accessory extremal

$$\xi = 0, \quad \tau = 0, \quad \sigma = 0, \; \kappa^* = 0, \tag{4.3.28}$$

leads to the necessary condition that, along the minimizing arc E of the primary Bolza problem,

$$\pi^T \begin{pmatrix} 0 & 0 & 0 \\ 0 & R_1 & R_2^T \\ 0 & R_2 & R_3 \end{pmatrix} \pi \geq 0, \qquad (4.3.29)$$

for all $(n+m) \times 1$ matrices π satisfying

$$(I\ B_1\ B_3)\pi = 0. \qquad (4.3.30)$$

Taking $\pi_{m+1}, \pi_{m+2}, \ldots, \pi_{n+m}$ to be arbitrary, condition (4.3.29) subjected to (4.3.30) is seen to imply that the $m \times n$ matrix

$$R_4 \equiv \begin{pmatrix} R_1 & R_2^T \\ R_2 & R_3 \end{pmatrix} \qquad (4.3.31)$$

must be positive semidefinite. Hence we have proved condition (ii). Q.E.D.

We will now prove condition (i). If Q_2B_2 is not symmetric we introduce new partition lines running between the m^{**} and $(m^{**}+1)$ rows/columns where $m^* < m^{**} \leq n-1$. Then

$$R \equiv \begin{pmatrix} R_1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \qquad (4.3.32)$$

$$Q \equiv \begin{pmatrix} Q_1 \\ Q_2^* \\ Q_2^{**} \end{pmatrix}, \text{ say,} \qquad (4.3.33)$$

$$B \equiv (B_1\ B_2^*\ B_2^{**}), \text{ say.} \qquad (4.3.34)$$

Furthermore m^{**} is chosen such that the $(m-m^{**}) \times (m-m^{**})$ order matrix $Q_2^{**}B_2^{**}$ is symmetric and is of the highest possible order. With m^{**} chosen in this manner, the non-symmetry of Q_2B_2 implies that

$$B_2^{**T}Q_2^{*T} - Q_2^{**}B_2^{*} \not\equiv 0. \tag{4.3.35}$$

This is easily seen on examining

$$Q_2B_2 \equiv \begin{pmatrix} Q_2^{*}B_2^{*} & Q_2^{*}B_2^{**} \\ Q_2^{**}B_2^{*} & Q_2^{**}B_2^{**} \end{pmatrix}. \tag{4.3.36}$$

If $B_2^{**T}Q_2^{*T} - Q_2^{**}B_2^{*} \equiv 0$, then the order of the symmetric matrix $Q_2^{**}B_2^{**}$ can be increased by at least one. This is because the $(m-m^{**}+1)\times(m-m^{**}+1)$ matrix consisting of $Q_2^{**}B_2^{**}$ and the last column of $Q_2^{**}B_2^{*}$ and the last row of $Q_2^{*}B_2^{**}$ and the last diagonal element of $Q_2^{*}B_2^{*}$ is symmetric. In the most extreme case $Q_2^{**}B_2^{**}$ is a one $\times$ one order matrix which is trivially symmetric. This accounts for $m^{**} \leqslant n-1$.

As $Q_2^{**}B_2^{**}$ is symmetric, condition (ii) of this theorem is applicable to the accessory minimum problem with matrices R, Q and B partitioned in the manner displayed in (4.3.32) to (4.3.34). Thus we are led to the condition that the matrix

$$R_4 \equiv \begin{pmatrix} R_1 & 0 & R_2^{*T} \\ 0 & 0 & R_2^{**T} \\ R_2^{*} & R_2^{**} & R_3^{*} \end{pmatrix}, \tag{4.3.37}$$

must be positive semidefinite, where

$$\begin{aligned}(R_2^{*}\ R_2^{**}) &\equiv B_2^{**T}(Q_1^{T}\ Q_2^{*T}) - Q_2^{**}(B_1\ B_2^{*}), \\ &= (B^{**T}Q_1^{T}-Q_2^{**}B_1 \,|\, B_2^{**T}Q_2^{*T}-Q_2^{**}B_2^{*}),\end{aligned} \tag{4.3.38}$$

Therefore,

$$R_2^* \equiv B_2^{**T} Q_1^T - Q_2^{**} B_1 , \tag{4.3.39}$$

$$R_2^{**} \equiv B_2^{**T} Q_2^{*T} - Q_2^{**} B_2^* . \tag{4.3.40}$$

The positive semidefiniteness of R_4 of (4.3.37) implies that the submatrix

$$\begin{pmatrix} 0 & R_2^{**T} \\ R_2^{**} & R_3^* \end{pmatrix} \tag{4.3.41}$$

must be positive semidefinite and by Lemma 1 we are led to the condition

$$R_2^{**} \equiv B_2^{**T} Q_2^{*T} - Q_2^{**} B_2^* \equiv 0 . \tag{4.3.42}$$

But from (4.3.35) this condition is not satisfied. Hence matrix $Q_2 B_2$ must be symmetric. Q.E.D.

<u>Corollary 2.1</u>.

Suppose that the control variable u_m appears linearly in (4.1.2) and (4.1.3). Let us equate to zero the variations of all the other control variables, that is, let

$$\dot{v}_r(t) = 0, \quad r = 1,2,\ldots,m-1. \tag{4.3.43}$$

Then the matrices R,Q,P,A,B of this reduced accessory minimum problem are

$$R = 0; \quad Q \equiv \left(-\frac{\partial^2 H}{\partial u_m \partial x_i}\right); \quad P \equiv \left(-\frac{\partial^2 H}{\partial x_i \partial x_j}\right); \tag{4.3.44}$$

$$A \equiv \left(\frac{\partial f_i}{\partial x_j}\right); \quad B \equiv \left(\frac{\partial f_i}{\partial u_m}\right) . \tag{4.3.45}$$

Assuming that this reduced accessory minimum problem remains normal, the method used to prove condition (ii) of Theorem 2 can be used to prove the necessary condition

$$B^TPB - \frac{d}{dt}(QB) - 2Q(AB - \dot{B}) \geq 0, \qquad (4.3.46)$$

along a singular minimizing arc E of the primary Bolza problem. Therefore we have

$$-\frac{\partial f_i}{\partial u_m}\frac{\partial^2 H}{\partial x_i \partial x_j}\frac{\partial f_j}{\partial u_m} + \frac{d}{dt}\left[\frac{\partial^2 H}{\partial u_m \partial x_i}\frac{\partial f_i}{\partial u_m}\right]$$
$$+ 2\frac{\partial^2 H}{\partial u_m \partial x_i}\left[\frac{\partial f_i}{\partial x_j}\frac{\partial f_j}{\partial u_m} - \frac{d}{dt}\left(\frac{\partial f_i}{\partial u_m}\right)\right] \geq 0 , \qquad (4.3.47)$$

where m is not summed and this condition is equivalent to the test first obtained by Kelley [19]. Note that Kelley used the other Hamiltonian function and this accounts for the difference in signs between the above result and that of Kelley.

<u>Corollary 2.2</u>.

If Q_2B_2 is identically symmetric and $R_2 \equiv 0$ and $R_3 \equiv 0$ along a singular extremal E then a new set of conditions is necessary at each element of E, namely:

(i) The $(m-m^*)\times(m-m^*)$ order matrix Q_3B_3 must be identically symmetric. The matrices Q_3 and B_3 are displayed in (4.3.22) and (4.3.8).

(ii) If Q_3B_3 is identically symmetric, the matrix

$$\begin{pmatrix} R_1 & R^T_{2,1} \\ R_{2,1} & R_{3,1} \end{pmatrix} , \tag{4.3.48}$$

must be positive semidefinite where

$$R_{2,1} \equiv B_3^T Q_1^T - Q_3 B_1 , \tag{4.3.49}$$

$$R_{3,1} \equiv B_3^T P B_3 - \frac{d}{dt}(Q_3 B_3) - Q_3 B_4 - B_4^T Q_3^T , \tag{4.3.50}$$

$$B_4 \equiv AB_3 - \dot{B}_3 . \tag{4.3.51}$$

In the same manner there exists a series of necessary conditions involving the symmetry of $Q_k B_k$ (k not summed, $k > 2$), and the positive semidefiniteness of

$$\begin{pmatrix} R_1 & R^T_{2,k-2} \\ R_{2,k-2} & R_{3,k-3} \end{pmatrix} \tag{4.3.52}$$

where

$$R_{2,k-2} \equiv B_k^T Q_1^T - Q_k B_1 , \tag{4.3.53}$$

$$R_{3,k-2} \equiv B_k^T P B_k - \frac{d}{dt}(Q_k B_k) - Q_k B_{k+1} - B_{k+1}^T Q_k^T . \tag{4.3.54}$$

$$Q_k \equiv B_{k-1}^T P - Q_{k-1} A - \dot{Q}_{k-1} , \tag{4.3.55}$$

$$B_{k+1} \equiv AB_k - \dot{B}_k , \tag{4.3.56}$$

assuming that for $h = 3,4,\ldots,k-1$, the matrices $R_{2,h-2}$ and $R_{3,h-2}$ vanish identically and $Q_h B_h$ is symmetric.

<u>Proof</u>

If $R_2 \equiv 0$ and $R_3 \equiv 0$ then from (4.3.27) we have

$$2\omega = \dot{\sigma}^T R_1 \dot{\sigma} + 2\dot{\sigma}^T Q_1 \tau + 2\dot{\kappa}^{*T} Q_3 \tau + \tau^T P \tau \tag{4.3.57}$$

and the equation of variation remains as

$$\dot{\tau} = A\tau + B_1\dot{\sigma} + B_3\dot{\kappa}^*. \tag{4.3.58}$$

These are in the standard forms displayed in (4.2.6) and (4.2.7) but with

$$R \equiv \begin{pmatrix} R_1 & 0 \\ 0 & 0 \end{pmatrix}, \tag{4.3.59}$$

$$Q \equiv \begin{pmatrix} Q_1 \\ Q_3 \end{pmatrix}, \tag{4.3.60}$$

$$B \equiv (B_1 \quad B_3). \tag{4.3.61}$$

Moreover the end variations appear in the transformed accessory minimum problem in the same manner as they appeared in the original accessory minimum problem. Hence this corollary may be proved by employing the same arguments as those used to prove theorem 2. Repeating in this manner the series of necessary conditions may be obtained.

Notes. 1. In corollary 2.2. we have limited ourselves to the special cases when the matrices R_2 and R_3 vanish identically. The more general case when only certain submatrices of R_2 and R_3 vanish identically could have been considered. However, this requires introducing many new matrices and we will lose sight of the main features of such singular problems.

2. Theorem 2 and corollaries remain valid but slightly modified when the control functions are subjected to the inequalities displayed in (4.1.6). The simplest method to

deal with such complications is to let the variation of any control variable function $u_r(t)$ attaining its bounds be zero. A more rigorous but tedious method is to employ the Valentine's device, Lawden [13], Berkovitz [27] and absorb the new multipliers into the matrix R_1.

3. In practice the series of necessary conditions in corollary 2.2. will either be effective or else it will indicate that a combination of the second and higher order variation of the functional J of (4.1.2) must be examined. "Effective" means that the conditions will either be violated or else strict inequalities are eventually imposed (except possibly at a finite number of points in the interval $[x_1, x_2]$).

4. On examining the matrices R_2, R_3 and B_3 displayed in (4.3.6) to (4.3.8) it is easily seen that the matrices B_2 and Q_2 occur most frequently. Therefore the calculations for the matrices R_2, R_3 and B_3 can be reduced if there are zero elements in B_2 and Q_2. From experience this is a common occurrence.

5. We will now make a note of other useful forms of the matrix R_3 of (4.3.7). Firstly, in obtaining (4.3.23) from (4.3.21) we have employed the quadratic form property

$$2\kappa^T Q_2 B_3 \kappa \equiv \kappa^T (Q_2 B_3 + B_3^T Q_2^T)\kappa. \qquad (4.3.62)$$

This enables R_3 of (4.3.7) to be a symmetric matrix. It is perfectly legitimate not to use this quadratic form property at this stage in which case we have instead of R_3 of (4.3.7),

$$B_2^T P B_2 - \frac{d}{dt} Q_2 B_2 - 2Q_2 B_3 \equiv R_3^* \text{ , say.} \tag{4.3.63}$$

By (4.3.8) this is again expressible as

$$\begin{aligned} R_2^* &\equiv B_2^T P B_2 - \dot{Q}_2 B_2 - Q_2 \dot{B}_2 - 2Q_2 A B_2 + 2Q_2 \dot{B}_2 , \\ &= B_2^T P B_2 + Q_2 \dot{B}_2 - \dot{Q}_2 B_2 - 2Q_2 A B_2 , \qquad (4.3.64) \\ &= (B_2^T P - Q_2 - Q_2 A) B_2 - Q_2 A B + Q_2 B_2 , \\ &= Q_3 B_2 - Q_2 B_3 , \qquad (4.3.65) \end{aligned}$$

by means of (4.3.22) and (4.3.8). This matrix R_3^* is however not necessarily symmetric.

Secondly, we have from (4.3.7),

$$\begin{aligned} R_3 &\equiv B_2^T P B_2 - \frac{d}{dt}(Q_2 B_2) - Q_2 B_3 - B_3^T Q_2^T , \\ &= B_2^T P B_2 - \dot{Q}_2 B_2 - Q_2 \dot{B}_2 - Q_2 A B_2 + Q_2 \dot{B}_2 - B_3^T Q_2^T , \\ &= (B_2^T P - \dot{Q}_2 - Q_2 A) B_2 - B_3^T Q_2^T , \\ &= Q_3 B_2 - B_2^T Q_3^T , \qquad (4.3.66) \end{aligned}$$

by means of (4.3.22).

4.4. VARIABLE THRUST ARCS FOR ROCKET FLIGHT IN A RESISTING MEDIUM

We will examine the extremal variable thrust arcs of a rocket moving in a resisting medium and in a vertical plane of a flat earth. This problem has been studied in detail by Miele [25]. We shall consider the simplified case when the thrust direction is always tangential to the velocity vector and when there are two degrees of freedom, namely, the lift program and the mass flow program. The case of one further degree of freedom, namely, the thrust direction program, becomes unmanageable analytically.

If X denotes a horizontal coordinate, h a vertical coordinate, V the magnitude of velocity, γ the angle between the velocity vector and horizontal direction, in the mass, g the acceleration of gravity, c the equivalent exit velocity of the rocket engine, β the mass flow, D the drag and L the lift, the equations of motion of the rocket are

$$\dot{X} = V \cos\gamma \ , \tag{4.4.1}$$

$$\dot{h} = V \sin\gamma \ , \tag{4.4.2}$$

$$\dot{V} = -g \sin\gamma - (D - c\beta)/m \ , \tag{4.4.3}$$

$$\dot{\gamma} = -(g \cos\gamma)/V + L/(mV), \tag{4.4.4}$$

$$\dot{m} = -\beta \tag{4.4.5}$$

and

$$0 \leqslant \beta \leqslant \beta_{max} \tag{4.4.6}$$

where

$$D = D(h,V,L) \quad \text{and} \quad g = \text{const.} \tag{4.4.7}$$

The problem is to minimize a certain terminal performance index $G(X,h,V,\gamma,m,t)$ subjected to conditions (4.4.1) to (4.4.6) and prescribed end conditions. As only extremal variable thrust arcs are being examined the constraint (4.4.6) is ignored. The control variables are L and β.

The Hamiltonian function is

$$H = \lambda_1 V \cos\gamma + \lambda_2 V \sin\gamma - \lambda_3[g \sin\gamma + (D - c\beta/m]$$
$$- \lambda_4[(g \cos\gamma)/V - L/(mV)] - \lambda_5\beta. \qquad (4.4.8)$$

The Euler-Lagrange equations are

$$\dot{\lambda}_1 = 0 , \qquad (4.4.9)$$
$$\dot{\lambda}_2 = (\lambda_3/m)(\partial D/\partial h), \qquad (4.4.10)$$
$$\dot{\lambda}_3 = -\lambda_1 \cos\gamma - \lambda_2 \sin\gamma + (\lambda_3/m)(\partial D/\partial V)$$
$$- (\lambda_4 g/V^2)\cos\gamma + \lambda_4 L/(mV^2), \qquad (4.4.11)$$
$$\dot{\lambda}_4 = \lambda_1 V \sin\gamma - \lambda_2 V \cos\gamma + \lambda_3 g \cos\gamma$$
$$- (\lambda_4 g/V)\sin\gamma , \qquad (4.4.12)$$
$$\dot{\lambda}_5 = -\lambda_3(D - c\beta)/m^2 + \lambda_4 L/(m^2 V), \qquad (4.4.13)$$
$$-(\lambda_3/m)\partial D/\partial L + \lambda_4/(mV) = 0 , \qquad (4.4.14)$$
$$(c\lambda_3/m) - \lambda_5 = 0 . \qquad (4.4.15)$$

Differentiating (4.4.15) we get

$$- c\dot{m}\lambda_3/m^2 + c\dot{\lambda}_3/m - \dot{\lambda}_5 = 0 \qquad (4.4.16)$$

Using equations (4.4.5), (4.4.11), (4.4.12) and (4.4.14) this equation gives

$$(\lambda_1\cos\gamma + \lambda_2\sin\gamma) = \frac{\lambda_3}{m}\left(\frac{D}{c} + \frac{\partial D}{\partial V}\right) - \frac{\lambda_4}{mV}\left[\left(\frac{1}{c} - \frac{1}{V}\right)L + \frac{mg}{V}\cos\gamma\right], \qquad (4.4.17)$$

$$= \frac{\lambda_3}{m}\left(\frac{D}{c} + \frac{\partial D}{\partial V}\right) - \frac{\lambda_3}{m}\,\frac{\partial D}{\partial L}\left[\left(\frac{1}{c} - \frac{1}{V}\right)L + \frac{mg}{V}\cos\gamma\right]. \qquad (4.4.18$$

We will show that the Euler-Lagrange equations and (4.4.18) imply that the matrices R_1, R_2, R_3 of theorem 2 are given by

$$R_1 = \left(-\frac{\partial^2 H}{\partial L^2}\right) = \frac{\lambda_3}{m}\frac{\partial^2 D}{\partial L^2}, \tag{4.4.19}$$

$$R_2 = \frac{c\lambda_3}{m^2}\left[\frac{\partial^2 D}{\partial L \partial V} + \frac{1}{V}\frac{\partial D}{\partial L}\right], \tag{4.4.20}$$

$$R_3 = \frac{c^2\lambda_3}{m^3}\left[\frac{\partial^2 D}{\partial V^2} + \frac{2}{c}\frac{\partial D}{\partial V} + \frac{D}{c^2}\right] + \frac{\lambda_3}{m^3 V^2}\frac{\partial D}{\partial L}[2mgc^2\cos\gamma + L(2cV - 2c^2 - V^2)]. \tag{4.4.21}$$

<u>Calculations</u>: The equations (4.4.5) and (4.4.14) will be frequently used to eliminate $\dot{m}$ and λ_4 respectively. We will use x_i to denote the set of state variables X, h, V, γ, m respectively. The symbols f_i will be used to denote the right hand side of the equations of motion. We have

$$B_2^T = \left(\frac{\partial f_i}{\partial \beta}\right) = \left(0, 0, \frac{c}{m}, 0, -1\right), \tag{4.4.22}$$

$$Q_2 = \left(-\frac{\partial^2 H}{\partial \beta \partial x_i}\right) = \left(0, 0, 0, 0, \frac{c\lambda_3}{m^2}\right). \tag{4.4.23}$$

It follows that

$$AB_2 = \left(\frac{\partial f_i}{\partial x_j}\right)\left(0, 0, \frac{c}{m}, 0, -1\right)^T$$

$$= \left(\frac{\partial f_i}{\partial V}\frac{c}{m} - \frac{\partial f_i}{\partial m}\right) = (C_i), \text{ say.} \tag{4.4.24}$$

$$\dot{B}_2^T = \left(0, 0, \frac{c\beta}{m^2}, 0, 0\right). \tag{4.4.25}$$

It follows that

$$\begin{aligned} -2Q_2B_3 &= -2Q_2C + 2Q_2\dot{B}_2, \\ &= -2Q_{25}C_5 + 2Q_{25}\dot{B}_{25}, \\ &= -\frac{c\lambda_3}{m^2}\left(\frac{\partial f_5}{\partial V}\frac{c}{m} - \frac{\partial f_5}{\partial m}\right) + 0, \\ &= 0. \end{aligned} \tag{4.4.26}$$

Now

$$Q_2B_2 = -\frac{c\lambda_3}{m^2}, \tag{4.4.27}$$

It follows that

$$\frac{d}{dt}Q_2B_2 = -\frac{c\dot{\lambda}_3}{m^2} - \frac{2c\lambda_3\beta}{m^3},$$

$$= \frac{c}{m^2}(\lambda_1\cos\gamma + \lambda_2\sin\gamma) - \frac{c\lambda_3}{m^3}\frac{\partial D}{\partial V} + \frac{c}{m^2}\frac{\lambda_4 g}{V^2}\cos\gamma - \frac{c\lambda_4}{m^3V^2}L - \frac{2c\lambda_3\beta}{m^3}. \tag{4.4.28}$$

$$R_3 = B_{23}^2P_{33} + 2B_{23}B_{25}P_{35} + B_{25}^2P_{55} - \frac{d}{dt}(Q_2B_2),$$

$$= -\frac{c^2}{m^2}\frac{\partial^2 H}{\partial V^2} + \frac{2c}{m}\frac{\partial^2 H}{\partial V\partial m} - \frac{\partial^2 H}{\partial m^2} - \frac{d}{dt}(Q_2B_2). \tag{4.4.29}$$

$$-\frac{c^2}{m^2}\frac{\partial^2 H}{\partial V^2} = \frac{c^2\lambda_3}{m^3}\frac{\partial^2 D}{\partial V^2} + \frac{2\lambda_4 gc^2}{m^2V^3}\cos\gamma - \frac{2c^2\lambda_4 L}{m^3V^3}, \tag{4.4.30}$$

$$\frac{2c}{m}\frac{\partial^2 H}{\partial V\partial m} = \frac{2c\lambda_3}{m^3}\frac{\partial D}{\partial V} + \frac{2c\lambda_4 L}{m^3V^2}, \tag{4.4.31}$$

$$-\frac{\partial^2 H}{\partial m^2} = \frac{2\lambda_3}{m^3}(D - c\beta) - \frac{2\lambda_4 L}{m^3V}. \tag{4.4.32}$$

It follows that

$$R_3 = -\frac{c}{m^2}(\lambda_1\cos\gamma + \lambda_2\sin\gamma) + \frac{c^2\lambda_3}{m^3}\left[\frac{\partial^2 D}{\partial V^2} + \frac{3}{c}\frac{\partial D}{\partial V} + \frac{2D}{c}\right] + \frac{\lambda_4}{mV}\left[\frac{2gc^2\cos\gamma}{mV^2} - \frac{cg\cos\gamma}{mV} + \frac{3cL}{m^2V} - \frac{2c^2L}{m^2V^2} - \frac{2L}{m^2}\right] \tag{4.4.33}$$

Then using (4.4.17) and (4.4.14) we get R_3 as displayed in (4.4.21).

$$R_2 = B_2^TQ_1^T - Q_2B_1,$$

$$= \frac{c}{m}Q_{13} - Q_{15} - \frac{c\lambda_3}{m^2}B_{15}. \tag{4.4.34}$$

$$Q_{13} = -\frac{\partial^2 H}{\partial L \partial V} = \frac{\lambda_3}{m}\frac{\partial^2 D}{\partial L \partial V} + \frac{\lambda_4}{mV^2} ,$$

$$= \frac{\lambda_3}{m}\frac{\partial^2 D}{\partial L \partial V} + \frac{\lambda_3}{mV}\frac{\partial D}{\partial L} . \qquad (4.4.35)$$

$$Q_{15} = -\frac{\partial^2 H}{\partial L \partial m} = -\frac{\lambda_3}{m^2}\frac{\partial D}{\partial L} + \frac{\lambda_4}{m^2 V} = 0 . \qquad (4.4.36)$$

$$B_{15} = \frac{\partial f_5}{\partial L} = 0 . \qquad (4.4.37)$$

It follows that

$$R_2 = \frac{c\lambda_3}{m^2}\left[\frac{\partial^2 D}{\partial L \partial V} + \frac{1}{V}\frac{\partial D}{\partial L}\right] ,$$

as displayed in (4.4.20). Q.E.D.

Theorem 2 implies that

$$R_1 \geqslant 0 , \qquad (4.4.38)$$

$$R_3 \geqslant 0 , \qquad (4.4.39)$$

$$R_1 R_3 - R_2^2 \geqslant 0 . \qquad (4.4.40)$$

The inequality condition (4.4.38) is contained in the classical Clebsch condition. From it we may deduce the sign of λ_3. The inequality (4.4.39) can also be obtained by Kelley's test [18]. Inequality (4.4.40) is a new optimality condition. Note that the matrix corresponding to $Q_2 B_2$ is trivially symmetric because it is a one $\times$ one order matrix. In general, if singularity is due to only one control variable appearing linearly, the matrix corresponding to $Q_2 B_2$ is trivially symmetric.

4.5. A CLASS OF IDENTICALLY SINGULAR OPTIMAL CONTROL PROBLEMS

Employing the indices displayed in (4.1.1), the statement of the problem is: Find control functions $u_r(t)$ which minimize the performance index

$$J \equiv g[x(t_1),t_1] + \int_{t_0}^{t_1} L_0(x,t) + u_r L_r(x,t)dt \ , \qquad (4.5.1)$$

with state variables $x_i(t)$ satisfying

$$\dot{x}_i = C_i(x,t) + D_{ir}(x,t)u_r \ , \qquad (4.5.2)$$

$$x_i(t_0) = x_{io} \ , \ \text{(constants)}, \qquad (4.5.3)$$

$$\psi^{\mu}[x(t_1),t_1] = 0 \ . \qquad (4.5.4)$$

and vector $(u_r) \in U$ which is an open region. Here L_o, L_r, C_i, D_{ir} are known functions of x_j and t. This problem is a generalization of the extremization of the general linear integral studied in the previous chapter.

The Euler-Lagrange equations are

$$\dot{\lambda}_i = - \lambda_j \frac{\partial C_j}{\partial x_i} - \lambda_j \frac{\partial D_{js}}{\partial x_i} u_s + \frac{\partial L_0}{\partial x_i} + u_s \frac{\partial L_s}{\partial x_i} \ , \qquad (4.5.5)$$

$$\lambda_i D_{ir} - L_r = 0. \qquad (4.5.6)$$

In the notation of theorem 2 the matrix R along a singular extremal E, satisfying these Euler-Lagrange equations, is zero and hence theorem 2 is applicable. The matrices

$$B_2 = \left(\frac{\partial f_i}{\partial u_r}\right) = (D_{ir}) \ , \qquad (4.5.7)$$

$$Q_2 = \left(- \frac{\partial^2 H}{\partial u_r \partial x_i}\right) = \left(-\lambda_j \frac{\partial D_{jr}}{\partial x_i} + \frac{\partial L_r}{\partial x_i}\right) \qquad (4.5.8)$$

It follows that

$$Q_2 B_2 = \left(-\lambda_j \frac{\partial D_{jr}}{\partial x_i} D_{is} + \frac{\partial L_r}{\partial x_i} D_{is}\right)$$

$$= (\pi_{rs}), \text{ say.} \qquad (4.5.9)$$

Condition (i) of theorem 2 requires that this $m \times m$ order matrix (π_{rs}) must be symmetric along the minimizing arc.

An example of such a problem was formulated by Breakwell [33] in connection with optimal guidance. The statement of the problem is: Find control variables $u(t)$ and $r(t)$ so as to minimize the performance index

$$J \equiv \int_0^T [2u\sqrt{p} + \alpha r]dt, \qquad (4.5.10)$$

with state variables $p(t)$, $g(t)$ satisfying

$$\dot{p} = -2\tau u p + r a g^2 \ , \qquad (4.5.11)$$

$$\dot{g} = - r a g^2 \ , \qquad (4.5.12)$$

where $\tau = T - t$ with T predetermined, α is a specified constant and $a(t)$ is a known function. We will assume that the end conditions are of the form (4.5.4) and that singular extremals involving intermediate levels of both $r(t)$ and $u(t)$ exist.

As displayed in (4.1.9) we define the Hamiltonian function

$$H \equiv \lambda(r a g^2 - 2ru p) - \mu r a g^2 - 2u\sqrt{p} - \alpha r. \qquad (4.5.13)$$

The Euler-Lagrange equations are

$$\dot{\lambda} = 2\lambda\tau u - u/\sqrt{p}, \qquad (4.5.14)$$

$$\dot{\mu} = 2(\mu - \lambda)r a g, \qquad (4.5.15)$$

$$-2\lambda\tau p - 2\sqrt{p} = 0 ,$$

$$\Rightarrow \quad \lambda\tau\sqrt{p} + 1 = 0 , \qquad (4.5.16)$$

$$\lambda a g^2 - \mu a g^2 - \alpha = 0 ,$$

$$\Rightarrow \quad \alpha + (\mu - \lambda) a g^2 = 0 . \qquad (4.5.17)$$

In the notation of theorem 2 we have

$$B_2 = \begin{pmatrix} -2\tau p & ag^2 \\ 0 & -ag^2 \end{pmatrix}, \qquad (4.5.18)$$

$$Q_2 = \begin{pmatrix} 2\lambda\tau+1/\sqrt{p} & 0 \\ 0 & -2(\lambda-\mu)a g \end{pmatrix},$$

$$= \begin{pmatrix} -1/\sqrt{p} & 0 \\ 0 & -2\alpha/g \end{pmatrix}, \qquad (4.5.19)$$

by means of (4.5.16) and (4.5.17). Hence

$$Q_2B_2 = \begin{pmatrix} 2\tau\sqrt{p} & -a g^2/\sqrt{p} \\ 0 & 2a\alpha g \end{pmatrix}. \qquad (4.5.20)$$

In order that Q_2B_2 is identically symmetric

$$a.g^2/\sqrt{p} \equiv 0 \ ,$$

$$\Rightarrow \qquad g \equiv 0 \ . \qquad (4.5.21)$$

This rules out the doubly singular extremals involving intermediate levels of both $r(t)$ and $u(t)$.

4.6. SINGULAR EXTREMALS IN LAWDEN'S PROBLEM

We will now apply theorem 2 to the problem of deciding the status of the intermediate-thrust arcs which arise in optimal rocket trajectory problems, Lawden [13], [34]. The conclusions to which we shall be led to, are in agreement with those obtained by Kopp and Moyer [19], Robbins [20], Kelley [16], Tait [22] and Gurley [21], employing different methods. In this section we will use the following set of indices: $i,j,k = 1,2,3$.

The optimal rocket trajectory problem may be formulated thus: $Ox_1x_2x_3$ is an inertial frame. At time t a rocket has coordinates x_i and velocity components v_i: its motor thrust acts in a direction having direction cosines ℓ_i and the mass rate of propellant consumption is m. Then if $g_i(x_1,x_2,x_3,t)$ are the components of the gravitational field in the frame, M is the rocket mass and c the exhaust velocity, the equations of motion are

$$\dot{v}_i = (c\,m\,\ell_i)/M + g_i \ , \qquad (4.6.1)$$

$$\dot{x}_i = v_i \qquad (4.6.2)$$

$$\dot{M} = - m. \tag{4.6.3}$$

Employing the usual spherical polar angles θ,ϕ the direction cosines can be expressed thus:

$$\ell_1 = \sin\theta \cos\phi, \quad \ell_2 = \sin\theta \sin\phi, \quad \ell_3 = \cos\theta. \tag{4.6.4}$$

The general problem is to transfer the rocket from a given set of initial conditions

$$v_i = v_{io}, \quad x_i = x_{io}, \quad M = M_o \quad \text{at} \quad t = t_o, \tag{4.6.5}$$

to a given set of terminal conditions

$$\psi^{\mu} [v_i(t_1), x_i(t_1), M(t_1), t_1] = 0, \quad (\mu = 1, 1, \ldots, N \leq 7), \tag{4.6.6}$$

such that a certain terminal function

$$J = J\,[v_i(t_1), x_i(t_1), M(t_1), t_1], \tag{4.6.7}$$

is minimized.

A specific example is the problem of achieving a docking rendezvous with an orbiting satellite when (4.6.6) becomes

$$v_i - v_i(t_1) = 0, \quad x_i - x_i(t_1) = 0, \tag{4.6.8}$$

where $v_i(t_1)$ and $x_i(t_1)$ are the equations of motion of the satellite. We may want the transit time to be minimized in which case

$$J = - t_1 . \tag{4.6.9}$$

A further equation of the type (4.6.6) may arise from fuel requirements. Alternatively we may want the propellant expenditure to be minimized in which case

$$J = - M_1 \ . \tag{4.6.10}$$

Another example is the problem of transferring the rocket to another given point at which it is to have a specific velocity and such that propellant expenditure is to be minimized. This leads to the terminal conditions

$$v_i = v_i \ , \quad x_i = x_i \ , \quad \text{at} \quad t = t_1 \tag{4.6.11}$$

and $$J = - M_1 .$$

In the now standard terminology the functions $v_i(t)$, $x_i(t)$, $M(t)$ are the state variables and $\theta(t)$, $\phi(t)$ and $m(t)$ are the control variables. From (4.1.9) we define the Hamiltonian function

$$H = \lambda_i(c\, m\, \ell_i/M + g_i) + \lambda_{i+3}v_i - \lambda_7 m . \tag{4.6.12}$$

Lawden [13],[34], has shown that there exists certain extremals known as the intermediate-thrust arcs (I - T arcs), which satisfy the multiplier rule and the Weierstrass condition. It may be verified that the determinant Δ of (4.2.14) vanishes identically along such extremals. This is because the control variable m occurs linearly in the Hamiltonian function.

The Euler-Lagrange equations are

$$\dot{\lambda}_i = - \lambda_{i+3} , \tag{4.6.13}$$

$$\dot{\lambda}_{i+3} = -\lambda_j \frac{\partial g_j}{\partial x_i} , \qquad (4.6.14)$$

$$\dot{\lambda}_7 = \frac{cm}{M^2} \lambda_i \ell_i , \qquad (4.6.15)$$

$$-\frac{cm}{M} \lambda_j \frac{\partial \ell_j}{\partial \theta} = 0 , \qquad (4.6.16)$$

$$-\frac{cm}{M} \lambda_j \frac{\partial \ell_j}{\partial \phi} = 0 , \qquad (4.6.17)$$

$$-c \frac{\lambda_i \ell_i}{M} + \lambda_7 = 0 . \qquad (4.6.18)$$

By means of (4.6.4), the equation (4.6.16) and (4.6.17) imply

$$\lambda_1/\ell_1 = \lambda_2/\ell_2 = \lambda_3/\ell_3 = p, \text{ say.} \qquad (4.6.19)$$

Differentiating (4.6.18) and employing (4.6.3) and (4.6.15) we have

$$\dot{\lambda}_7 = \frac{c}{M} \frac{d}{dt} (\lambda_i \ell_i) - \frac{c\dot{M}}{M^2} \lambda_i \ell_i$$

$$\Rightarrow \lambda_i \ell_i = \text{const} = p , \qquad (4.6.20)$$

by means of (4.6.19) and $\ell_i \ell_i = 1$.

Denoting the variations of the state variables v_i, x_i, M by $\eta_i, \eta_{3+i}, \eta_7$ respectively and the variations of the control variables θ, ϕ, m by $\zeta_8, \zeta_9, \zeta_{10}$ respectively, the equations of of variation prove to be

$$\dot{\eta}_i = \frac{cm}{M} \left(\frac{\partial \ell_i}{\partial \theta} \zeta_8 + \frac{\partial \ell_i}{\partial \phi} \zeta_9 \right) + \frac{1}{M} c \ell_i \zeta_{10} + \frac{\partial g_i}{\partial x_j} \eta_{3+j} - \frac{cm}{M^2} \ell_i \eta_7 , \qquad (4.6.21)$$

$$\dot{\eta}_{3+i} = \eta_i , \qquad (4.6.22)$$

$$\dot{\eta}_7 = -\zeta_{10} . \qquad (4.6.23)$$

The integrand of the second variation is

$$2\omega = \frac{pcm}{M}\,(\zeta_8^2 + \zeta_9^2\sin^2\theta) + \frac{2pc}{M^2}\zeta_{10}\eta_7 - \lambda_k \frac{\partial^2 g_k}{\partial x_i \partial x_j}\,\eta_{3+i}\eta_{3+j} - \frac{2mpc}{M^3}\,\eta_7^2. \tag{4.6.24}$$

By employing (4.6.23) this can be simplified thus:

$$\int_{t_0}^{t_1} \frac{2pc}{M^2}\,\zeta_{10}\eta_7 dt = -\int_{t_0}^{t_1} \frac{pc}{M^2}\,\frac{d}{dt}\,(\eta_7^2)\,dt$$

$$= -\left[\frac{pc}{M^2}\,\eta_7^2\right]_{t_0}^{t_1} + \int_{t_0}^{t_1} \frac{2mpc}{M^3}\,\eta_7^2 dt, \tag{4.6.25}$$

where use has been made of (4.6.3) after integration by parts. Absorbing the terms involving end values in the form 2γ, the integrand 2ω reduces to the form

$$2\omega = \frac{pcm}{M}\left(\zeta_8^2 + \zeta_9^2\sin^2\theta\right) - \lambda_k \frac{\partial^2 g_k}{\partial x_i \partial x_j}\,\eta_{3+i}\eta_{3+j}. \tag{4.6.26}$$

It is clear that 2ω and the equations of variation are in the forms (4.2.6) and (4.2.7). The motivation of the above simplification is to reduce the matrix Q_2, in the notation of theorem 2, to zero. This reduces considerably the necessary calculations. Again we have

$$B_2^T \equiv (c\ell_i/M, 0, 0, 0, -1), \tag{4.6.27}$$

$$Q_2 = 0\ , \tag{4.6.38}$$

$$R_1 = \begin{pmatrix} pcm/M & 0 \\ 0 & (pcm/M)\sin^2\theta \end{pmatrix}, \tag{4.6.29}$$

$$Q_1 = 0 \ , \tag{4.6.30}$$

$$B_1^T = \begin{pmatrix} \frac{cm}{M}\frac{\partial \ell_i}{\partial\theta}, & 0, & 0, & 0, & 0 \\ \frac{cm}{M}\frac{\partial \ell_i}{\partial\phi}, & 0, & 0, & 0, & 0 \end{pmatrix}, \tag{4.6.31}$$

$$A = \begin{pmatrix} 0 & \frac{\partial g_i}{\partial x_j} & -\frac{cm\ell_i}{M^2} \\ \delta_{ij} & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \tag{4.6.32}$$

$$P = \begin{pmatrix} 0 & 0 & 0 \\ 0 & -\lambda_k \frac{\partial^2 g_k}{\partial x_i \partial x_j} & 0 \\ 0 & 0 & 0 \end{pmatrix}. \tag{4.6.33}$$

The orders of the matrices B_1, B_2 and A are 7×2, 7×1 and 7×7 respectively.

From (4.3.6) and (4.3.7) we have

$$R_2 = B_2^T.0 - 0.B_1 = 0, \tag{4.6.34}$$

$$R_3 = B_2^T P B_2 = 0. \tag{4.6.35}$$

Therefore it is necessary to apply corollary 2.2. From (4.3.55) and (4.3.56) we have

$$Q_3 = B_2^T P - Q_2 A - \dot{Q}_2 = 0 \ , \tag{4.6.36}$$

$$B_3^T = (AB_2 - \dot{B}_2)^T$$

$$= \left(\frac{cm\ell_i}{M^2} - \frac{d}{dt}\frac{c\ell_i}{M}, \ \frac{c\ell_i}{M}, \ 0\right). \tag{4.6.37}$$

From (4.3.53) and (4.3.54) we get

$$R_{2,1} = B_3^T.0 - 0.B_1 = 0 \,, \tag{4.6.38}$$

$$R_{3,1} = B_3^T P B_3 = -\frac{c^2}{M^2}\, \ell_i \ell_j \lambda_k \frac{\partial^2 g_k}{\partial x_i \partial x_j} \,. \tag{4.6.39}$$

The positive semidefiniteness of the matrix displayed in (4.3.52) requires that

$$\frac{pcm}{M} \geqslant 0, \ \frac{pcm}{M} \sin^2\theta \geqslant 0, \ \left(\frac{c}{M}\right)^2 \ell_i \ell_j \lambda_k \frac{\partial^2 g_k}{\partial x_i \partial x_j} \leqslant 0. \tag{4.6.40}$$

It follows that p must be positive and together with (4.6.19) we have

$$\lambda_i \lambda_j \lambda_k \frac{\partial^2 g_k}{\partial x_i \partial x_j} \leqslant 0. \tag{4.6.41}$$

In the special case when the gravitational field is an inverse square law of attraction towards the origin, we have

$$g_k = -\frac{\mu x_k}{r^3} \,, \tag{4.6.42}$$

where $r^2 = x_i x_i$, and it follows that

$$\frac{\partial^2 g_k}{\partial x_i \partial x_j} = \frac{3\mu}{r^5}\left(\delta_{ij} x_k + \delta_{jk} x_i + \delta_{ki} x_j\right) - \frac{15 x_i}{r^7}\, x_j x_k. \tag{4.6.43}$$

Hence,

$$\lambda_i \lambda_j \lambda_k \frac{\partial^2 g_k}{\partial x_i \partial x_j} = \frac{9\mu}{r^5}\left(\lambda_i \lambda_i \lambda_j x_j\right) - \frac{15\mu}{r^7}\,(\lambda_i x_i)^3. \tag{4.6.44}$$

But $\lambda_i \lambda_i = p^2$ and $\lambda_i x_i = p\, r \cos\psi$, where ψ is the angle between the radius vector from O to the rocket and the direction of thrust. Condition (4.6.41) is accordingly

$$\frac{3\mu p^3 s}{r^4}\,(3 - 5s^2) \leqslant 0, \tag{4.6.45}$$

where $s = \cos\psi$ and hence

$$\text{either} \qquad s \geqslant (3/5)^{\frac{1}{2}} \quad \text{or} \quad 0 \geqslant s \geqslant -(3/5)^{\frac{1}{2}}. \tag{4.6.46}$$

We will now show that all the I-T arcs satisfy the further inequality

$$s^2 \leqslant 1/3 \ . \tag{4.6.47}$$

This requires differentiating twice the condition of singularity

$$\lambda_i\lambda_i = p^2 = \text{const.} \tag{4.6.48}$$

Employing (4.6.13) and (4.6.14) we have

$$\lambda_i\dot{\lambda}_i = 0 = -\,\lambda_i\lambda_{3+i} \ , \tag{4.6.49}$$

and
$$\dot{\lambda}_i\lambda_{3+i} + \lambda_i\dot{\lambda}_{3+i} = 0 \ ,$$

$$\Rightarrow \lambda_{3+i}\lambda_{3+i} + \lambda_i\lambda_j\,\frac{\partial g_j}{\partial x_i} = 0 \ . \tag{4.6.50}$$

With g_i displayed in (4.6.42),

$$\lambda_i\lambda_j\,\frac{\partial g_j}{\partial x_i} = -\,\frac{\mu p^2}{r^3} + \frac{3\mu p^2}{r^3}\cos^2\psi \tag{4.6.51}$$

and as λ_{3+i} is real, $\lambda_{3+i}\lambda_{3+i}$ must be non-negative. Hence from (4.6.50) and (4.6.51) we get

$$3\mu p^2\cos^2\psi - \mu p^2 \leqslant 0$$

$$\Rightarrow \qquad 3s^2 \leqslant 1. \quad \text{Q.E.D.}$$

The inequalities (4.6.46) and (4.6.48) rule out those I-T arcs satisfying the condition

$$0 \leqslant s \leqslant (1/3)^{\frac{1}{2}} . \tag{4.6.52}$$

Lawden has shown that all the two-dimensional I-T arcs where the transit time is not predetermined satisfy the inequality (4.6.52). Hence they are non-optimal. The I-T arcs that require further examination are those satisfying the second inequality of (4.6.46). Along such arcs the thrust direction are directed towards the centre of attraction. At the present moment it is felt that such arcs occur in the two-dimensional case where the transit time is given and in the three dimensional problem.

4.7. SINGULAR CONTROL FOR LINEAR SYSTEMS

In this section the indices (4.1.1) will be employed. The statement of the problem is: Find control functions $u_r(t)$ which minimizes the performance index

$$J \equiv g[x(t_1),t_1] + \int_{t_0}^{t_1} u^TRu + 2u^TQx + x^TPxdt, \tag{4.7.1}$$

with state variables $x_i(t)$ satisfying

$$\dot{x}_i = A_{ij}(t)x_j + B_{ir}(t)u_r, \tag{4.7.2}$$

$$x_i(t_o) = x_{io}, \text{ (constants)} , \tag{4.7.3}$$

$$\psi^{\mu}[x(t_1),t_1] = 0 \tag{4.7.4}$$

and vector $(u_r) \in U$, which is an open region. Moreover

$$R(t) = \begin{pmatrix} R_1 & 0 \\ 0 & 0 \end{pmatrix}, \tag{4.7.5}$$

where R_1 is a $m^* \times m^*$ positive definite matrix. Matrices Q, B and u are partitioned similarly:

$$Q(t) = \begin{pmatrix} Q_1 \\ Q_2 \end{pmatrix} , \quad B = (B_1 \ B_2), \quad u = \begin{pmatrix} v \\ w \end{pmatrix} . \tag{4.7.6}$$

In order that theorem 2 is satisfied we assume the further hypotheses:

(i) Q_2B_2 is identically symmetric,

(ii) the matrix

$$R_4(t) = \begin{pmatrix} R_1 & R_2^T \\ R_2 & R_3 \end{pmatrix}, \tag{4.7.7}$$

is positive definite, where

$$R_2(t) = B_2^T Q_1^T - Q_2B_1, \tag{4.7.8}$$

$$R_3(t) = B_2^T P B_2 - \frac{d}{dt}(Q_2B_2) - B_3^T Q_2^T - Q_2B_3, \tag{4.7.9}$$

and

$$B_3 = AB_2 - \dot{B}_2 . \tag{4.7.10}$$

This class of problems is important because the dynamical equations of many engineering systems are linear.

The motivation of this section is to demonstrate the existence of singular extremals and to lay down a procedure for their derivation. In matrix notation the scalar Hamiltonian function

$$H(\lambda,x,u,t) = \lambda^T Ax + \lambda^T Bu - v^T R_1 v - 2u^T Qx - x^T Px. \quad (4.7.11)$$

The Euler-Lagrange equation is

$$\begin{aligned}\dot{\lambda} &= -A^T\lambda + 2Q^T u + 2Px \\ &= -A^T\lambda + 2Q_1^T v + 2Q_2^T w + 2Px\ ,\end{aligned} \quad (4.7.12)$$

$$B_1^T\lambda - 2R_1 v - 2Q_1 x = 0, \quad (4.7.13)$$

$$B_2^T\lambda - 2Q_2 x = 0. \quad (4.7.14)$$

Also it will be useful to write (4.7.2) as

$$\dot{x} = Ax + B_1 v + B_2 w. \quad (4.7.15)$$

Differentiating (4.7.14) we have

$$\dot{B}_2^T\lambda + B_2^T\dot{\lambda} - 2\dot{Q}_2 x - 2Q_2\dot{x} = 0 \quad (4.7.16)$$

and employing (4.7.12) and (4.7.15) we get

$$\begin{aligned}\dot{B}_2^T\lambda - B_2^T A^T\lambda + 2B_2^T Q_1^T v + 2B_2^T Q_2^T w + 2B_2^T Px - 2\dot{Q}_2 x \\ - 2Q_2 Ax - 2Q_2 B_1 v - 2Q_2 B_2 w = 0.\end{aligned} \quad (4.7.17)$$

By hypothesis Q_2B_2 is symmetric. Hence

$$2B_2^TQ_2^Tw - 2Q_2B_2w = 0 \; . \tag{4.7.18}$$

Therefore (4.7.17) becomes

$$2(B_2^TQ_1^T-Q_2B_1)v + (\dot{B}_2^T-B_2^TA^T)\lambda + 2(B_2^TP-Q_2A-\dot{Q}_2)x = 0.$$

Introducing the notation

$$Q_3 = B_2^TP - Q_2A - \dot{Q}_2 \tag{4.7.19}$$

and employing (4.7.8) and (4.7.10) we have

$$2R_2v - B_3^T\lambda + 2Q_3x = 0 \; . \tag{4.7.20}$$

Differentiating, we are led to

$$2\dot{R}_2v + 2R_2\dot{v} - \dot{B}_3^T\lambda - B_3^T\dot{\lambda} + 2\dot{Q}_3x + 2Q_3\dot{x} = 0$$

and by means of (4.7.12) and (4.7.15) this equation becomes

$$\begin{aligned}2\dot{R}_2v - 2B_3^TQ_1^Tv + 2Q_3B_1v + 2R_2\dot{v} + 2Q_3B_2w - 2B_3^TQ_2^Tw - \dot{B}_3^T\lambda + B_3^TA^T\lambda \\ - 2B_3^TPx + 2\dot{Q}_3x + 2Q_3Ax = 0,\end{aligned}$$

It follows that

$$\begin{aligned}2(Q_3B_2 - B_3^TQ_2^T)w + 2R_2\dot{v} + 2(Q_3B_1 - B_3^TQ_1^T + \dot{R}_2)v \\ - B_4^T\lambda - Q_4x = 0,\end{aligned} \tag{4.7.21}$$

where

$$B_4 = AB_3 - \dot{B}_3 \; , \tag{4.7.22}$$

$$Q_4 = B_3^TP - Q_3A - \dot{Q}_3. \tag{4.7.23}$$

From (4.3.66) the matrix R_3 of (4.7.9) may also be written thus

$$R_3 = Q_3B_2 - B_3^TQ_2^T . \tag{4.7.24}$$

By hypothesis, R_4 of (4.7.7) is positive definite. Hence R_3 is positive definite and consequently it is nonsingular. Therefore (4.7.21) may be solved for the vector w in terms of v, $\dot{v}$, λ, x, t. Unless $R_2 = 0$ this solution of w invariably contains $\dot{v}$.

A more complete approach is to eliminate $\dot{v}$ by means of (4.7.13). Differentiating (4.7.13) we have

$$\dot{B}_1^T\lambda + B_1^T\dot{\lambda} - 2\dot{R}_1v - 2R_1\dot{v} - 2\dot{Q}_1x - 2Q_1\dot{x} = 0. \tag{4.7.25}$$

Employing (4.7.12) and (4.7.15) this equation becomes

$$2R_1\dot{v} = (\dot{B}_1^T - B_1^TA^T)\lambda + 2(B_1^TQ_2^T - Q_1B_2)w + 2(B_1^TQ_1^T - Q_1B_1 - \dot{R}_1)v$$
$$+ 2(B_1^TP - Q_1A - \dot{Q}_1)x = 0,$$

It follows that

$$2\dot{v} = R_1^{-1}(\dot{B}_1^T - B_1^TA^T)\lambda - 2R_1^{-1}R_2^Tw + 2R_1^{-1}(B_1^TQ_1^T - Q_1B_1 - \dot{R}_1)v$$
$$+ 2R_1^{-1}(B_1^TP - Q_1A - \dot{Q}_1)x = 0. \tag{4.7.26}$$

Substituting this into (4.7.21) and by (4.7.24) we have

$$2R_3w - 2R_2R_1^{-1}R_2^Tw + 2R_2R_1^{-1}(\dot{B}_1^T - B_1^TA^T)\lambda - B_4^T\lambda$$
$$+ 2R_2R_1^{-1}(B_1^TQ_1^T - Q_1B_1 - \dot{R}_1)v + 2(Q_3B_1 - B_3^TQ_1^T + \dot{R}_2)v - Q_4x$$
$$+ 2R_2R_1^{-1}(B_1^TP - Q_1A - \dot{Q}_1)x = 0. \tag{4.7.27}$$

The coefficient of w is

$$2(R_3 - R_2R_1^{-1}R_2^T). \tag{4.7.28}$$

From Gantamacher [31, p.46], the determinant of R_4 of (4.7.7) may be expressed as

$$|R_4| = |R_1||R_3 - R_2R_1^{-1}R_2^T|. \tag{4.7.29}$$

Since R_4 and R_1 are positive definite and consequently non-singular therefore we have

$$|R_1||R_3 - R_2R_1^{-1}R_2^T| > 0 \tag{4.7.30}$$

$$\Rightarrow \quad |R_3 - R_2R_1^{-1}R_2^T| > 0. \tag{4.7.31}$$

Hence (4.7.27) can be solved for w in terms of λ, x, v, t.

The procedure for the derivation of the singular extremals is now clear. Firstly (4.7.13) is solved for the vector v in terms of λ, x, t. This is possible as R_1 is nonsingular. Then this vector v is substituted into the solution w of (4.7.27). The vector functions v and w of λ, x, t are then substituted into (4.7.2) and (4.7.12) and we are led to a system of 2n first order differential equations in λ, x, t, which in principle can be solved. Subsequently the imposition of the constraint (4.7.14) reduces the dimension of the solution space to a lower order.

4.8. ON THE DERIVATION OF SINGULAR EXTREMALS

In this section we will examine how singular extremals may be derived. It will be assumed that the independent

variable t does not occur explicitly in (4.1.2) to (4.1.5)
If otherwise, t is eliminated by means of $x_{n+1} = t$ and the equation

$$\dot{x}_{n+1} = 1 \ , \tag{4.8.1}$$

is adjoined to (4.1.3) and x_{n+1} is considered as a state variable.

In the special case in which the problem involves only a single control variable appearing linearly in the Hamiltonian function it may be shown, Tait [22], that corollary 2.2. is equivalent to

$$(-1)^m \frac{\partial}{\partial u}\left[\frac{d^{2m}}{dt^{2m}}\left(\frac{\partial H}{\partial u}\right)\right] \leqslant 0, \tag{4.8.2}$$

with $$\frac{d^k}{dt^k}\left(\frac{\partial H}{\partial u}\right) = 0 \quad \text{for} \quad k = 1,2,\ldots,2m. \tag{4.8.3}$$

Note that our Hamiltonian is defined by (4.1.9) whereas that of Tait and Kelley is defined by (4.1.8). If the inequality (4.8.2) is satisfied strictly it means that the equation

$$\frac{d^{2m}}{dt^{2m}}\left(\frac{\partial H}{\partial u}\right) = 0 \ , \tag{4.8.4}$$

can be solved for the control variable u in terms of λ_i and x_i.

The procedure for deriving the singular extremals is then clear. It consists of solving (4.8.4) for the control variable u as a function of λ_i and x_i and substituting it into the equations

$$\dot{x}_i = f_i(x,u) \tag{4.8.5}$$

and

$$\dot{\lambda}_i = -\frac{\partial H}{\partial x_i}(\lambda,x,u) \;. \tag{4.8.6}$$

The general solution of this system of differential equations contains 2n arbitrary constants. However, (4.8.3) must be satisfied and the final solution space is of a lower dimension, which depends on how many of (4.8.3) are independent equations.

We will now generalize these remarks to the problem when the control variables u_r for $r = m^* + 1,\ m^* + 2, \ldots, m$ appear linearly in (4.1.2) and (4.1.3). The indices (4.1.1) will be employed. The notation of theorem 2 will be used. We will assume that the matrix Q_2B_2 is identically symmetric and R_4 of (4.3.5) is positive definite. We will firstly show that

$$R_1 = \left(-\frac{\partial^2 H}{\partial u_\alpha \partial u_\beta}\right) , \tag{4.8.7}$$

$$R_2 = \left(\frac{\partial}{\partial u_\beta}\frac{d}{dt}\frac{\partial H}{\partial u_\rho}\right) , \tag{4.8.8}$$

$$R_3 - R_2R_1^{-1}R_2^T = \left(\frac{\partial}{\partial u_\nu}\left[\frac{d^2}{dt^2}\left(\frac{\partial H}{\partial u_\rho}\right)\right]\right) . \tag{4.8.9}$$

<u>Calculations</u>: With the Hamiltonian function $H(\lambda,x,u)$ defined by (4.1.9) the matrix R_1 is obtained by definition. The positive definiteness of R_4 implies that R_1 is positive

definite and is consequently nonsingular.

It is convenient to rewrite

$$\dot{x}_i = f_i(x,u) \quad \text{as} \quad \dot{x}_i = \frac{\partial H}{\partial \lambda_i} \tag{4.8.10}$$

The Euler-Lagrange equations are

$$\dot{\lambda}_i = -\frac{\partial H}{\partial x_i} \ , \tag{4.8.11}$$

$$\frac{\partial H}{\partial u_\alpha} = 0 \ , \tag{4.8.12}$$

and

$$\frac{\partial H}{\partial u_\rho} = 0 \ . \tag{4.8.13}$$

We have, by definition,

$$A = (\partial^2 H/\partial \lambda_i \partial x_j) \ , \tag{4.8.14}$$

$$B_1 = (\partial^2 H/\partial \lambda_i \partial u_\alpha) \ , \tag{4.8.15}$$

$$B_2 = (\partial^2 H/\partial \lambda_i \partial u_\rho) \ , \tag{4.8.16}$$

$$Q_1 = (-\partial^2 H/\partial u_\alpha \partial x_i) \ , \tag{4.8.17}$$

$$Q_2 = (-\partial^2 H/\partial u_\rho \partial x_i) \ , \tag{4.8.18}$$

$$P = (-\partial^2 H/\partial x_i \partial x_j) \ . \tag{4.8.19}$$

It follows that

$$R_2 = B_2^T Q_1^T - Q_2 B_1 = (R_{2,\rho\beta}) \ ,$$

$$= \left(-\frac{\partial^2 H}{\partial \lambda_i \partial u_\rho} \frac{\partial^2 H}{\partial u_\beta \partial x_i} + \frac{\partial^2 H}{\partial u_\rho \partial x_i} \frac{\partial^2 H}{\partial \lambda_i \partial u_\beta} \right) . \tag{4.8.20}$$

$$B_3 = AB_2 - \dot{B}_2 = (B_{3,i\rho}) \ ,$$

$$= \left(\frac{\partial^2 H}{\partial x_j \partial \lambda_i} \frac{\partial^2 H}{\partial \lambda_j \partial u_\rho} - \frac{d}{dt} \frac{\partial^2 H}{\partial \lambda_i \partial u_\rho} \right) . \tag{4.8.21}$$

$$Q_3 = B_2^T P - Q_2 A - \dot{Q}_2 = (Q_{3,\rho i})$$
$$= \left(- \frac{\partial^2 H}{\partial\lambda_j \partial u_\rho} \frac{\partial^2 H}{\partial x_j \partial x_i} + \frac{\partial^2 H}{\partial u_\rho \partial x_j} \frac{\partial^2 H}{\partial\lambda_j \partial x_i} + \frac{d}{dt} \frac{\partial^2 H}{\partial u_\rho \partial x_i} \right) . \qquad (4.8.22)$$

From (4.3.66) we have

$$R_3 = Q_3 B_2 - B_3^T Q_2^T = (R_{3,\rho\nu})$$
$$= \left(- \frac{\partial^2 H}{\partial\lambda_j \partial u_\rho} \frac{\partial^2 H}{\partial x_j \partial x_i} \frac{\partial^2 H}{\partial\lambda_i \partial u_\nu} + \frac{\partial^2 H}{\partial u_\rho \partial x_j} \frac{\partial^2 H}{\partial\lambda_j \partial x_i} \frac{\partial^2 H}{\partial\lambda_i \partial u_\nu} \right.$$
$$+ \frac{d}{dt} \left[\frac{\partial^2 H}{\partial u_\rho \partial x_i} \right] \frac{\partial^2 H}{\partial\lambda_i \partial u_\nu} + \frac{\partial^2 H}{\partial\lambda_j \partial u_\rho} \frac{\partial^2 H}{\partial x_j \partial\lambda_i} \frac{\partial^2 H}{\partial u_\nu \partial x_i}$$
$$\left. - \frac{d}{dt} \left[\frac{\partial^2 H}{\partial\lambda_i \partial u_\rho} \right] \frac{\partial^2 H}{\partial u_\nu \partial x_i} \right) . \qquad (4.8.23)$$

We will now consider the consequences of differentiating (4.8.13). We recall, section 4.2., that the assumption that the control variables u_ρ appear linearly in H and the further assumption that the classical Clebsch condition is satisfied imply

$$\frac{\partial^2 H}{\partial u_\rho \partial u_\alpha} \equiv 0 \; , \quad \frac{\partial^2 H}{\partial u_\rho \partial u_\nu} \equiv 0 \; . \qquad (4.8.24)$$

It is useful to note that H is linear in λ_i.

Employing (4.8.10) and (4.8.11) we have

$$\frac{d}{dt} \frac{\partial H}{\partial u_\rho} = - \frac{\partial^2 H}{\partial\lambda_i \partial u_\rho} \frac{\partial H}{\partial x_i} + \frac{\partial^2 H}{\partial u_\rho \partial u_i} \frac{\partial H}{\partial\lambda_i} = 0 \; . \qquad (4.8.25)$$

It follows that

$$\left(\frac{\partial}{\partial u_\nu}\frac{d}{dt}\frac{\partial H}{\partial u_\rho}\right) = \left(-\frac{\partial^2 H}{\partial \lambda_i \partial u_\rho}\frac{\partial^2 H}{\partial u_\nu \partial x_i} + \frac{\partial^2 H}{\partial u_\rho \partial x_i}\frac{\partial^2 H}{\partial \lambda_i \partial u_\nu}\right) ,$$
$$= (B_2^T Q_2^T - Q_2 B_2) \; ,$$
$$= 0 \; , \tag{4.8.26}$$

as Q_2B_2 is assumed to be identically symmetric. Again, we have

$$\left(\frac{\partial}{\partial u_\beta}\frac{d}{dt}\frac{\partial H}{\partial u_\rho}\right) = \left(-\frac{\partial^2 H}{\partial \lambda_i \partial u_\rho}\frac{\partial^2 H}{\partial u_\beta \partial x_i} + \frac{\partial^2 H}{\partial u_\rho \partial x_i}\frac{\partial^2 H}{\partial \lambda_i \partial u_\beta}\right)$$
$$= R_2 \tag{4.8.27}$$

from (4.8.20). Therefore we have proved (4.8.8). The other partial derivatives of (4.8.25) are

$$\frac{\partial}{\partial x_i}\frac{d}{dt}\frac{\partial H}{\partial u_\rho} = -\frac{\partial^2 H}{\partial \lambda_j \partial u_\rho}\frac{\partial^2 H}{\partial x_j \partial x_i} + \frac{\partial^2 H}{\partial u_\rho \partial x_j}\frac{\partial^2 H}{\partial \lambda_j \partial x_i}$$
$$- \frac{\partial^3 H}{\partial \lambda_j \partial u_\rho \partial x_i}\frac{\partial H}{\partial x_j} + \frac{\partial^3 H}{\partial u_\rho \partial x_j \partial x_i}\frac{\partial H}{\partial \lambda_j} , \tag{4.8.28}$$

$$\frac{\partial}{\partial \lambda_i}\frac{d}{dt}\frac{\partial H}{\partial u_\rho} = -\frac{\partial^2 H}{\partial u_\rho \partial \lambda_j}\frac{\partial^2 H}{\partial x_j \partial \lambda_i} \quad \frac{\partial^3 H}{\partial u_\rho \partial x_j \partial \lambda_i}\frac{\partial H}{\partial \lambda_j} . \tag{4.8.29}$$

Next we note that

$$\frac{d}{dt}\frac{\partial^2 H}{\partial u_\rho \partial x_i} = -\frac{\partial^3 H}{\partial u_\rho \partial x_i \partial \lambda_j}\frac{\partial H}{\partial x_j} + \frac{\partial^3 H}{\partial u_\rho \partial x_i \partial x_j}\frac{\partial H}{\partial \lambda_j} , \tag{4.8.30}$$

$$\frac{d}{dt}\frac{\partial^2 H}{\partial \lambda_i \partial u_\rho} = \frac{\partial^3 H}{\partial \lambda_i \partial u_\rho \partial x_j}\frac{\partial H}{\partial \lambda_j} . \tag{4.8.31}$$

Employing these, (4.8.28) and (4.8.29) become

$$\frac{\partial}{\partial x_i}\frac{d}{dt}\frac{\partial H}{\partial u_\rho} = -\frac{\partial^2 H}{\partial \lambda_j \partial u_\rho}\frac{\partial^2 H}{\partial x_j \partial x_i} + \frac{\partial^2 H}{\partial u_\rho \partial x_j}\frac{\partial^2 H}{\partial \lambda_j \partial x_i} + \frac{d}{dt}\frac{\partial^2 H}{\partial u_\rho \partial x_i} , \tag{4.8.32}$$

$$\frac{\partial}{\partial\lambda_i}\frac{d}{dt}\frac{\partial H}{\partial u_\rho} = -\frac{\partial^2 H}{\partial u_\rho \partial\lambda_j}\frac{\partial^2 H}{\partial x_j \partial\lambda_i} + \frac{d}{dt}\frac{\partial^2 H}{\partial\lambda_i \partial u_\rho} \, . \qquad (4.8.33)$$

We will now examine the second time-derivative of (4.8.13): we have

$$\begin{aligned}\frac{d^2}{dt^2}\left(\frac{\partial H}{\partial u_\rho}\right) &= \frac{d}{dt}\left[\frac{d}{dt}\left(\frac{\partial H}{\partial u_\rho}\right)\right] \\ &= \dot{u}_\alpha \frac{\partial}{\partial u_\alpha}\frac{d}{dt}\frac{\partial H}{\partial u_\rho} - \frac{\partial}{\partial\lambda_i}\left[\frac{d}{dt}\frac{\partial H}{\partial u_\rho}\right]\frac{\partial H}{\partial x_i} \\ &\quad + \frac{\partial}{\partial x_i}\left[\frac{d}{dt}\frac{\partial H}{\partial u_\rho}\right]\frac{\partial H}{\partial\lambda_i} \\ &= 0 \, . \qquad (4.8.34)\end{aligned}$$

Employing (4.8.26), we have

$$\begin{aligned}\frac{\partial}{\partial u_\nu}\left[\frac{d^2}{dt^2}\left(\frac{\partial H}{\partial u_\rho}\right)\right] &= \frac{\partial}{\partial u_\nu}\left[\dot{u}_\alpha \frac{\partial}{\partial u_\alpha}\left(\frac{d}{dt}\frac{\partial H}{\partial u_\rho}\right)\right] - \frac{\partial}{\partial\lambda_i}\left[\frac{d}{dt}\frac{\partial H}{\partial u_\rho}\right]\frac{\partial^2 H}{\partial u_\nu \partial x_i} \\ &\quad + \frac{\partial}{\partial x_i}\left[\frac{d}{dt}\frac{\partial H}{\partial u_\rho}\right]\frac{\partial^2 H}{\partial\lambda_i \partial u_\nu} \, , \qquad (4.8.35)\end{aligned}$$

Again employing (4.8.26), we have

$$\begin{aligned}\frac{\partial}{\partial u_\nu}\left[\dot{u}_\alpha \frac{\partial}{\partial u_\alpha}\frac{d}{dt}\frac{\partial H}{\partial u_\rho}\right] &= \frac{\partial \dot{u}_\alpha}{\partial u_\nu}\frac{\partial}{\partial u_\alpha}\frac{d}{dt}\frac{\partial H}{\partial u_\rho} + \dot{u}_\alpha \frac{\partial}{\partial u_\alpha}\frac{\partial}{\partial u_\nu}\frac{d}{dt}\frac{\partial H}{\partial u_\rho} \\ &= \frac{\partial \dot{u}_\alpha}{\partial u_\nu}\frac{\partial}{\partial u_\alpha}\left[\frac{d}{dt}\frac{\partial H}{\partial u_\rho}\right] . \qquad (4.8.36)\end{aligned}$$

To evaluate $\partial\dot{u}_\alpha/\partial u_\nu$, we differentiate (4.8.12), which gives

$$\frac{\partial^2 H}{\partial u_\beta \partial u_\alpha}\dot{u}_\alpha - \frac{\partial^2 H}{\partial u_\beta \partial\lambda_i}\frac{\partial H}{\partial x_i} + \frac{\partial^2 H}{\partial u_\beta \partial x_i}\frac{\partial H}{\partial\lambda_i} = 0. \qquad (4.8.37)$$

As matrix R_1 is nonsingular this equation can be solved for

$\dot{u}_\alpha$. Hence

$$\dot{u}_\alpha = R^{-1}_{1,\beta\alpha}\left[-\frac{\partial^2 H}{\partial u_\beta \partial \lambda_i}\frac{\partial H}{\partial x_i} + \frac{\partial^2 H}{\partial u_\beta \partial x_i}\frac{\partial H}{\partial \lambda_i}\right], \tag{4.8.38}$$

It follows that

$$\begin{aligned}\left(\frac{\partial \dot{u}_\alpha}{\partial u_\nu}\right) &= R_1^{-1}\left[-\frac{\partial^2 H}{\partial u_\beta \partial \lambda_i}\frac{\partial^2 H}{\partial x_i \partial u_\nu} + \frac{\partial^2 H}{\partial u_\beta \partial x_i}\frac{\partial^2 H}{\partial \lambda_i \partial u_\nu}\right] \\ &= R_1^{-1}[B_1^T Q_2^T - Q_1 B_2] \\ &= -\, R_1^{-1} R_2^T \,. \end{aligned} \tag{4.8.39}$$

We note that R_1, B_1 and Q_1 are independent of u_ν because of (4.8.24). Therefore from (4.8.36), (4.8.39) and (4.8.8) we have

$$\begin{aligned}\left(\frac{\partial}{\partial u_\nu}\left[\dot{u}_\alpha \frac{\partial}{\partial u_\alpha}\frac{d}{dt}\frac{\partial H}{\partial u_\rho}\right]\right) &= \left(\frac{\partial}{\partial u_\alpha}\frac{d}{dt}\frac{\partial H}{\partial u_\rho}\right)(-R_1^{-1}R_2^T) \\ &= -\, R_2 R_1^{-1} R_2^T \,. \end{aligned} \tag{4.8.40}$$

Employing (4.8.32) and (4.8.33) the other terms of (4.8.35) are

$$\begin{aligned} &-\frac{\partial}{\partial \lambda_i}\left[\frac{d}{dt}\frac{\partial H}{\partial u_\rho}\right]\frac{\partial^2 H}{\partial u_\nu \partial x_i} + \frac{\partial}{\partial x_i}\left[\frac{d}{dt}\frac{\partial H}{\partial u_\rho}\right]\frac{\partial^2 H}{\partial \lambda_i \partial u_\nu} \\ &\quad = +\frac{\partial^2 H}{\partial \lambda_j \partial u_\rho}\frac{\partial^2 H}{\partial x_j \partial \lambda_i}\frac{\partial^2 H}{\partial u_\nu \partial x_i} - \frac{d}{dt}\left[\frac{\partial^2 H}{\partial \lambda_i \partial u_\rho}\right]\frac{\partial^2 H}{\partial u_\nu \partial x_i} \\ &\qquad - \frac{\partial^2 H}{\partial \lambda_j \partial u_\rho}\frac{\partial^2 H}{\partial x_j \partial x_i}\frac{\partial^2 H}{\partial \lambda_i \partial u_\nu} + \frac{\partial^2 H}{\partial u_\rho \partial x_j}\frac{\partial^2 H}{\partial \lambda_j \partial x_i}\frac{\partial^2 H}{\partial \lambda_i \partial u_\nu} \\ &\qquad + \frac{d}{dt}\left[\frac{\partial^2 H}{\partial u_\rho \partial x_i}\right]\frac{\partial^2 H}{\partial \lambda_i \partial u_\nu} \\ &\quad = R_{3,\rho\nu} \,, \end{aligned} \tag{4.8.41}$$

from (4.8.23). Therefore (4.8.35), (3.8.40) and (4.8.41) imply

$$\left(\frac{\partial}{\partial u_\nu}\left[\frac{d^2}{dt^2}\,\frac{\partial H}{\partial u_\rho}\right]\right) = R_3 - R_2R_1^{-1}R_2^T,$$

which is what we set out to prove, see (4.8.9).

We recall that the matrix R_4, in the notation of theorem 2, is assumed to be positive definite. Therefore R_1 is nonsingular. From Gantamacher [31,p.46]

$$\begin{aligned} |R_4| &= |R_1|\,|R_3 - R_2R_1^{-1}R_2^T| \\ &> 0 \,. \end{aligned} \tag{4.8.42}$$

Hence $(R_3 - R_2R_1^{-1}R_2^T)$ is nonsingular.

The procedure for deriving the singular extremals is thus:

(1) Differentiate (4.8.12) with respect to time and solve for $\dot{u}_\alpha(\lambda,x)$. From (4.8.37) this is possible as R_1 is nonsingular.

(2) Compute the equations

$$\frac{d^2}{dt^2}\,\frac{\partial H}{\partial u_\rho} = 0 \,, \tag{4.8.43}$$

which will involve $\dot{u}_\alpha, u_\alpha, u_\rho, \lambda_i, x_i$. The variables $\dot{u}_\alpha$ occur only linearly and are eliminated by means of $\dot{u}_\alpha(\lambda,x)$ obtained in the previous step. We are thus led to an equation involving only $u_\alpha, u_\rho, \lambda_i, x_i$.

(3) As $(R_3 - R_2R_1^{-1}R_2^T)$ is nonsingular, by (4.8.9) this equation

can be solved for u_ν in terms of u_α, λ_i, x_i.

(4) The equation (4.8.12) is solved for u_α in terms of λ_i and x_i. This solution $u_\alpha(\lambda,x)$ does not contain u_ρ explicitly as $\partial^2 H/\partial u_\alpha \partial u_\rho \equiv 0$. Substituting into $u_\rho(u_\alpha,\lambda,x)$ of the previous step we obtain solutions $u_r(\lambda,x)$.

(5) These functions $u_r(\lambda,x)$ are substituted into (4.8.10) and (4.8.11). The general solution of this system of differential equations in λ_i and x_i contains 2n arbitrary constants. However, the imposition of the conditions

$$Q_2 B_2 = B_2^T Q_2^T \tag{4.8.44}$$

and

$$\frac{\partial H}{\partial u_\rho} = 0, \; \frac{d}{dt}\frac{\partial H}{\partial u_\rho} = 0, \tag{4.8.45}$$

reduces the number of these arbitrary constants.

CHAPTER 5. SINGULAR EXTREMALS IN THE GENERAL BOLZA PROBLEM.

5.1. A PRELIMINARY TRANSFORMATION.

We will now consider the accessory minimum problem of the Bolza problem as formulated in section 1.1.1. We will assume that the integers m, m^*, n satisfy the inequalities

$$0 \leqslant m \leqslant m^* < n. \tag{5.1.1}$$

The problem is supposed to be singular, that is, the determinant Δ of (1.3.9) vanishes identically along the reference minimizing arc E.

Employing matrix notation, the integrand of the second variation can be written thus:

$$2\omega = \eta'^T R \eta' + 2\eta'^T Q \eta + \eta^T P \eta, \tag{5.1.2}$$

and the m differential constraints (1.3.10), become

$$\phi\eta' + \theta\eta = 0. \tag{5.1.3}$$

We will show that we can always reduce the accessory minimum problem into a form in which the matrices R, Q, ϕ have the forms

$$R = \begin{pmatrix} 0 & 0 \\ 0 & R^* \end{pmatrix}, \tag{5.1.4}$$

$$Q = \begin{pmatrix} 0 \\ \overline{Q^*} \end{pmatrix}, \tag{5.1.5}$$

$$\phi = (I_n \mid -B^*), \tag{5.1.6}$$

where partition lines run between the mth and (m+1)th rows/columns and I_m is the $m \times m$ identity matrix.

In general, the matrices R, Q, ϕ of (5.1.2) and (5.1.3) are of the form

$$R = \begin{pmatrix} R_2 & R_3^T \\ R_3 & R_1 \end{pmatrix}, \tag{5.1.7}$$

$$Q = \begin{pmatrix} \underline{Q_1} \\ Q_2 \end{pmatrix}, \tag{5.1.8}$$

$$\phi = (N \mid M), \tag{5.1.9}$$

where partition lines run between the mth and (m+1)th rows/columns. According to the hypothesis of section 1.2, ϕ must be of rank m everywhere in Γ_1. Hence we can arrange the nomenclature so that N is nonsingular. Thus N^{-1} exists and premultiplying (5.1.3) by this matrix, we obtain

$$(I_m \mid A)\eta' + B\eta = 0, \tag{5.1.10}$$

where

$$A = N^{-1}M, \tag{5.1.11}$$

$$B = N^{-1}\theta. \tag{5.1.12}$$

Partitioning η thus:

$$\eta^T = (\rho^T \mid \pi^T), \tag{5.1.13}$$

where ρ is a $m \times 1$ matrix and π is a $(n-m) \times 1$ matrix, we have, employing (5.1.7) and (5.1.8),

$$\eta'^T R\eta' = \rho'^T R_2\rho' + 2\pi'^T R_3\rho' + \pi'^T R_1\pi' , \qquad (5.1.14)$$

$$2\eta'^T Q\eta = 2\rho'^T Q_1\eta + 2\pi'^T Q_2\eta. \qquad (5.1.15)$$

Equation (5.1.10) may be written as

$$\rho' + A\pi' + B\eta = 0. \qquad (5.1.16)$$

Eliminating ρ' from (5.1.14) and (5.1.15) by means of (5.1.16) we find that

$$\eta'^T R\eta' = \pi'^T R^*\pi' + 2\pi'^T K\eta + \eta^T L\eta, \qquad (5.1.17)$$

$$2\eta'^T Q\eta = 2\pi'^T G\eta + \eta^T H\eta, \qquad (5.1.18)$$

where

$$L = B^T L_2 B , \qquad (5.1.19)$$

$$K = A^T R_2 B - R_3 B , \qquad (5.1.20)$$

$$R^* = A^T R_2 A - R_3 A - A^T R_3^T + R_1 , \qquad (5.1.21)$$

$$G = Q_2 - A^T Q_1 \qquad (5.1.22)$$

$$H = -BQ_1 - Q_1^T B^T . \qquad (5.1.23)$$

It is convenient to introduce the notation

$$Q^* = G + K + Q_2 , \qquad (5.1.24)$$

$$P^* = H + L + P, \qquad (5.1.25)$$

$$B^* = -A, \; A^* = -B. \qquad (5.1.26)$$

Clearly R^* and P^* are symmetric.

Substituting (5.1.17) and (5.1.18) into (5.1.2) we are led to

$$2\omega = \pi'^T R^*\pi' + 2\pi'^T Q^*\eta + \eta^T P^*\eta. \qquad (5.1.27)$$

Employing (5.1.26), (5.1.16) becomes

$$\rho' - B^*\pi' - A^*\eta = 0. \qquad (5.1.28)$$

Comparing (5.1.27) with (5.1.2) and (5.1.28) with (5.1.3) the matrices R, Q, ϕ are in the forms displayed in (5.1.4) to (5.1.6).

We will now prove that the reduced form of the accessory minimum problem with matrices R, Q, ϕ displayed in (5.1.4) to (5.1.6), is singular if and only if the original accessory minimum problem is singular. We define an $(n+m)\times(n+m)$ matrix C in the partitioned form

$$C = \begin{pmatrix} D & -A & 0 \\ 0 & I_{n-m} & 0 \\ 0 & 0 & I_m \end{pmatrix} , \qquad (5.1.29)$$

where $D = N^{-1}$. Taking determinants and expanding by minors in the lower n rows we find that

$$|C| = |D| = |N|^{-1} \neq 0. \qquad (5.1.30)$$

Thus, C is regular. Also

$$C^T \begin{pmatrix} R & \phi^T \\ \phi & 0 \end{pmatrix} C = C^T \begin{pmatrix} R_2 & R_3^T & N^T \\ R_3 & R_1 & M^T \\ N & M & 0 \end{pmatrix} C,$$

$$= \begin{pmatrix} E_1 & E_2 & E_3 \\ E_4 & E_5 & E_6 \\ E_7 & E_8 & 0 \end{pmatrix}, \qquad (5.1.31)$$

where $E_1 = D^T R_2 D,$ (5.1.32)

$$E_2 = D^T R_3^T - D^T R_2 A, \tag{5.1.33}$$

$$E_3 = D^T N^T = (ND)^T = I_m, \tag{5.1.34}$$

$$E_4 = (R_3 - A^T R_2)D = E_2^T, \tag{5.1.35}$$

$$E_5 = A^T R_2 A - A^T R_3^T - R_3 A + R_1 = R^*, \tag{5.1.36}$$

$$\begin{aligned} E_6 &= -A^T N^T + M^T = (-NA + M)^T \\ &= (-NN^{-1}M + M)^T = 0, \end{aligned} \tag{5.1.37}$$

$$E_7 = ND = I_m, \tag{5.1.38}$$

$$E_8 = M - NA = 0, \tag{5.1.39}$$

by (5.1.11). Thus (5.1.31) can be written

$$C^T \begin{pmatrix} R & \phi^T \\ \phi & 0 \end{pmatrix} C = \begin{pmatrix} E_1 & E_4^T & I_m \\ E_4 & R^* & 0 \\ I_m & 0 & 0 \end{pmatrix}. \tag{5.1.40}$$

Taking determinants of both members of this equation and expanding the right-hand determinant by the last m rows followed by the last m columns we find that

$$\begin{aligned} |N|^{-2}\Delta &= (-1)^{(n+2)m + (n-m+2)m}|R^*| \\ &= (-1)^m |R^*|. \end{aligned} \tag{5.1.41}$$

It follows that $|R^*|$ vanishes if, and only if, Δ of (1.3.9) vanishes, that is if the minimizing arc is singular.

But the reduced accessory minimum problem is singular if the determinant

$$\begin{vmatrix} 0 & 0 & I_m \\ 0 & R^* & -B^{*T} \\ I_m & -B^* & 0 \end{vmatrix} = (-1)^m |R^*|, \qquad (5.1.42)$$

vanishes identically and we have accordingly proved that the reduced accessory minimum problem is singular if, and only if, the original accessory minimum problem is singular.

5.2. THE GENERALIZED CLEBSCH CONDITION.

In this section we will not deduce the generalized Clebsch condition in an explicit form but lay down a procedure by which it can be deduced in any particular problem. The procedure consists of three steps in the following sequence:

(i) reduction of the accessory minimum problem into the form where the matrices R, Q, ϕ are in the forms displayed in (5.1.4) to (5.1.6),

(ii) translation of the reduced accessory minimum problem into the state-control variables formulation,

(iii) application of theorem 2 to the transformed accessory minimum problem.

The first step has already been discussed in the previous section. The second step is carried out by adjoining to (5.1.28)

$$\pi' = u, \qquad (5.2.1)$$

where u is defined to be the (n - m) control vector. The n order state vector is given by

$$\eta^T = (\rho^T \mid \pi^T). \tag{5.2.2}$$

Eliminating π' from (5.1.27) and (5.1.28) by means of (5.2.1) we have

$$2\omega = u^T R^* u + 2u^T Q^* \eta + \eta^T P^* \eta, \tag{5.2.3}$$

and $$\rho' = A^* \eta + B^* u, \tag{5.2.4}$$

which together with (5.2.1) becomes

$$\eta' = \begin{pmatrix} A^* \\ 0 \end{pmatrix} \eta + \begin{pmatrix} B^* \\ I_{n-m} \end{pmatrix} u, \tag{5.2.5}$$

which is in the form (4.2.7) and 2ω of (5.2.3) is in the form (4.2.6). Hence theorem 2 and corollaries are applicable.

We will demonstrate this procedure by applying it to the problem studied in Chapter 3 above. This will then be an alternative proof of theorem 1. For convenience, we will employ a notation different from that of theorem 1 above. In this problem there are no side constraints in the primary variational problem and hence the equations of variation of the type (5.1.3) are absent. Step (i) of the procedure is therefore redundant.

From (3.2.4), (3.2.11) and (3.3.1) to (3.3.3) we have

$$2\omega = \eta'^T R \eta' + 2\eta'^T Q \eta + \eta^T P \eta, \tag{5.2.6}$$

where $$R = \begin{pmatrix} R_{11} & 0 \\ 0 & 0 \end{pmatrix}, \tag{5.2.7}$$

$$Q = \begin{pmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{pmatrix}, \tag{5.2.8}$$

$$P = \begin{pmatrix} P_{11} & P_{21}^T \\ P_{21} & P_{22} \end{pmatrix}, \tag{5.2.9}$$

$$\eta^T = (\sigma^T \mid \kappa^T) , \tag{5.2.10}$$

with partition lines running between the mth and (m+1)th rows/columns, $(0 \leqslant m < n)$.

Step (ii) consists of defining the n order control vector thus:

$$\eta' = u, \tag{5.2.11}$$

which is then used to eliminate η' from (5.2.6). We are thus led to

$$2\omega = u^T Ru + 2u^T Q\eta + \eta^T P\eta. \tag{5.2.12}$$

Before applying theorem 2 we first note that in its notation

$$B = I_n, \; B_1 = \begin{pmatrix} I_m \\ 0 \end{pmatrix}, \; B_2 = \begin{pmatrix} 0 \\ I_{n-m} \end{pmatrix}, \tag{5.2.13}$$

$$Q = Q, \; Q_1 = (Q_{11} \mid Q_{12}), \; Q_2 = (Q_{21} \mid Q_{22}), \tag{5.2.14}$$

$$R = R, \; R_1 = R_{11}, \tag{5.2.15}$$

$$P = P = \begin{pmatrix} P_{11} & P_{21}^T \\ P_{21} & P_{22} \end{pmatrix}, \tag{5.2.16}$$

$$A = 0. \tag{5.2.17}$$

Hence the application of theorem 2 to this accessory minimum problem leads us to conclude that along a singular minimizing arc:

(i) $$Q_2B_2 = (Q_{21} \mid Q_{22})\begin{pmatrix} 0 \\ I_{n-m} \end{pmatrix} = Q_{22}, \tag{5.2.18}$$

must be identically symmetric,

(ii) if Q_{22} is identically symmetric, the matrix

$$R_4 = \begin{pmatrix} R_1 & R_2^T \\ R_2 & R_3 \end{pmatrix}, \tag{5.2.19}$$

must be positive semidefinite where

$$R_2 = (0 \mid I_{n-m})\begin{pmatrix} Q_{11}^T \\ Q_{12}^T \end{pmatrix} - (Q_{21} \mid Q_{22})\begin{pmatrix} I_m \\ 0 \end{pmatrix}$$

$$= Q_{12}^T - Q_{21}, \tag{5.2.20}$$

$$R_3 = B_2^T P B_2 - (d/dx)(Q_2B_2) - Q_2B_3 - B_3^T Q_2^T,$$

and since

$$B_3 = AB_2 - B_2' = 0 - 0 = 0, \tag{5.2.21}$$

$$B_2^T P B_2 = (0 \mid I_{n-m})\begin{pmatrix} P_{11} & P_{21}^T \\ P_{21} & P_{22} \end{pmatrix}\begin{pmatrix} 0 \\ \hline I_{n-m} \end{pmatrix}$$

$$= P_{22}, \tag{5.2.22}$$

$$Q_2B_2 = Q_{22}, \text{ from (5.2.18)},$$

hence $$R_3 = P_{22} - Q_{22}' . \tag{5.2.23}$$

Thus, we have proved theorem 1.

We will now prove corollary 1.1. Suppose

$$R_2 \equiv 0, \quad R_3 \equiv 0, \tag{5.2.24}$$

where R_2 and R_3 are displayed in (5.2.20) and (5.2.23). We will now apply corollary 2.2 to this accessory minimum problem. From (4.3.22) we have

$$Q_3 = (0 \mid I_{n-m}) \begin{pmatrix} P_{11} & P_{21}^T \\ P_{21} & P_{22} \end{pmatrix} - (Q_{21}' \quad Q_{22}')$$

$$= (P_{21} - Q_{21}' \mid P_{22} - Q_{22}') . \tag{5.2.25}$$

With B_3 given by (5.2.21),

$$Q_3 B_3 = Q_3 \,.\, 0 = 0 \,, \tag{5.2.26}$$

which is identically symmetric. From (4.3.49) to (4.3.51) we have

$$\begin{aligned} R_{2,1} &= 0 \quad - Q_3 B_1 \\ &= P_{21} - Q_{21}' \,, \end{aligned} \tag{5.2.27}$$

$$R_{3,1} = 0 \,, \tag{5.2.28}$$

because $B_3 = 0$ from (5.2.21). Therefore we conclude that

$$\begin{pmatrix} R_{11} & P_{21}^T - Q_{21}'^T \\ P_{21} - Q_{21}' & 0 \end{pmatrix} , \tag{5.2.29}$$

must be positive semidefinite. By Lemma 1, we deduce that

$$P_{21} - Q_{21}' \equiv 0. \tag{5.2.30}$$

Hence we have proved corollary 1.1.

REFERENCES

[1] G.A. BLISS, The evolution of problems of the calculus of variations, Amer. Math. Monthly, 43 (1936) pp. 598-609.

[2] G.A. BLISS, "Lectures on the Calculus of Variations", Univ. of Chicago Press, Chicago 1946.

[3] O. BOLZA, "Vorlesungen uber Variationrechnung", Teubner 1909, reprinted by Chelsea Publ. Co., New York 1964.

[4] W.T. REID, A matrix differential equation of Riccati type, Amer. J. of Maths. 68 (1946) pp. 237-246.

[5] G.A. BLISS, Jacobi's condition for problems of the calculus of variations in parametric form, Trans. Amer. Math. Soc. 17 (1916) pp. 195-206.

[6] D.M. SMITH, Jacobi's condition for the problem of Lagrange in the calculus of variations, Trans. Amer. Math. Soc., 17 (1916) pp. 459-475.

[7] G.A. BLISS, Some recent developments in the calculus of variations, Bull. Amer. Math. Soc. 26 (1919) pp. 343-361.

[8] M. MORSE and W. LEIGHTON, Singular quadratic functionals, Trans. Amer. Math. Soc. 40 (1936) pp.252-286.

[9] O. BOLZA, "Lectures on the Calculus of Variations", Chicago 1904, reprinted by Dover Publ. Inc., New York 1961.

[10] J.D. MANCILL, Identically non-regular problems in the calculus of variations, Revista Matematica Y Fisica Teorica, Ser. A, 7 (1950) pp. 131-139.

[11] A. MIELE, Extremization of linear integrals by Green's theorem, Chap. 3 of "Optimization Techniques", (G. Leitmann, ed.) Academic Press, New York, 1962.

[12] G.W. HAYNES, On the optimality of a totally singular vector control: An extension of the Green's theorem approach to higher dimensions, Final Report, Contract NAS-2-2351, Martin Company, Denver, Colorado, U.S.A. 1965.

[13] D.F. LAWDEN, "Optimal Trajectories for Space Navigation", Butterworths Inc., Washington, D.C. 1963.

[14] G. LEITMANN, Rocket trajectory optimization, Applied Mech. Rev., 18 (1965) pp. 345-350.

[15] B. PAIEWONSKY, Optimal control: A review of theory and practice, AIAA J., 3 (1965) pp. 1985-2006.

[16] H.J. KELLEY, R.E. KOPP and H.G. MOYER, Singular Extremals, Chap. 3 of "Optimization - Theory and Applications: Vol. 1 A Variational Approach." (G. Leitmann, ed.) Academic Press, New York, to be published.

[17] C.D. JOHNSON, Singular solutions in problems of optimal control, Chap. 4 of "Advances in Control Systems", Vol. 2. (C.T. Leondes, ed.), Academic Press, New York, 1965.

[18] H.J. KELLEY, A second variation test for singular extremals, AIAA J., 2 (1964) pp. 1380-1382.

[19] R.E. KOPP and H.G. MOYER, Necessary conditions for singular extremals, AIAA J., 3, (1965) pp. 1439-1444.

[20] H.M. ROBBINS, Optimality of intermediate-thrust arcs of rocket trajectories, AIAA J., 3, (1965) pp. 1095-1098.

[21] J.G. GURLEY, Optimal thrust trajectories in an arbitrary gravitational field, J. SIAM, CONTROL, Ser. A. 2 (1964) pp. 423-432.

[22] K. TAIT, Singular problems in optimal control, Ph.D. Thesis, Havard Univ., Camb., Mass., U.S.A. (1965).

[23] B.S. GOH, The second variation of the singular Bolza problem, J. SIAM, CONTROL, Ser. A. Vol. 4, (1966) p.p. 309-325.

[24] B.S. GOH, Necessary conditions for singular extremals involving multiple control variables, J. SIAM, CONTROL, Ser. A. Vol. 4, to be published.

[25] A. MIELE, The calculus of variations in applied aero-dynamics and flight mechanics, Chapter 4 of "Optimization Techniques" (G. Leitmann ed.), Academic Press, New York, 1962.

[26] M.R. HESTENES, Elements of the calculus of variations, Chapter 4 of "Modern Mathematics for the Engineer" (E.F. Beckenback ed.), Univ. of California Extension Series, McGraw-Hill Bk. Co. Inc., New York, 1956.

[27] L.D. BERKOVITZ, Variational methods of control and programming, J. of Math. Anal. and Appl., 3(1961) pp. 145-169.

[28] L.A. PARS, "An Introduction to the Calculus of Variations". Heinemann, London, 1962.

[29] H. HANCOCK, "Theory of Maxima and Minima" Boston 1917, reprinted by Dover Publ. Inc., New York, 1960.

[30] T.W. CHAUNDY, Maxima and minima: The semi-definite case, The Quarterly J. of Maths., 4 (1933) pp. 1-11.

[31] E.F. GANTMACHER, "Theory of Matrices," Vol. 1, Chapter 2, Chelsea Publ. Co., New York 1959.

[32] H.D. BLOCK, On the minimality of variational principles of classical particle mechanics, Amer. Math. Monthly, 62 (1955) pp. 161 - 168.

[33] J.V. BREAKWELL, A doubly singular problem in optimal interplanetary guidance, J. SIAM, CONTROL, Ser. A, 3 (1965) pp. 71 - 77.

ADDITIONAL REFERENCES

[1] J.V. BREAKWELL and F.F. TUNG, An optimal information and control policy in interplanetary guidance, Third International Federation of Automatic Control Congress, London, June 20 - 25, 1966.

[2] J.L. KELLER, On minimum propellant paths for thrust limited rockets, Astronautica Acta, 10 (1964) pp. 262 - 269.

[3] R.A. ROHRER and M. SOBRAL Jr., Optimal singular solutions for linear multi-input systems, J. of Basic Engineering, Trans. ASME, Ser. D 88 (1966) pp. 323 - 328.