SECOND ORDER ENDPOINT CONDITION

By Richard G. Brusch and Thomas L. Vincent

Prepared under Grant No. NGR 03-002-011 by
THE UNIVERSITY OF ARIZONA
Tucson, Arizona

for

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION

March 1970

PRECEDING PAGE BLANK NOT FILMED.

TABLE OF CONTENTS

TABLE OF CONTENTS - Cont'd

LIST OF ILLUSTRATIONS

ABSTRACT

The nature of multiple solutions to a general Problem of Bolza in the calculus of variations is investigated. These multiple stationary solutions are of several distinct types, each of which is briefly discussed.

Multiple stationary solutions often arise when only first order necessary conditions have been applied. A second order test to eliminate some multiple solutions of this type is developed. It is shown that second order tests can be broken down into path tests and endpoint tests and that once the first order necessary conditions of the calculus of variations have been applied, the second order path test and the endpoint test can be applied independently.

Only the second order endpoint test is investigated. Direct application of this condition requires the analytical integration of a set of nonlinear differential equations subject to mixed boundary conditions. Since this is often difficult or impossible, an algorithm is developed for numerically implementing the second order endpoint condition. A geodetics problem is solved analytically to illustrate the theory and demonstrates that the second order condition is an effective computational tool, by eliminating certain classes of non-optimal solutions from consideration.

INTRODUCTION

1.1 The Problem of Bolza

Since much of what is to follow depends upon an understanding of the Problem of Bolza as formulated in control notation (Vincent and Brusch, 1966, pp. 4-5), a brief statement of the problem in its simplest form is appropriate.

Among the set of all continuous state functions,

$$y_i(t) \qquad i = 1, 2, \ldots, n; \; t_o \leqq t \leqq t_f \tag{1.1.1}$$

and continuous control variable functions

$$u_k(t) \qquad k = 1, 2, \ldots, m < n \tag{1.1.2}$$

satisfying differential equations and end-conditions of the form

$$\dot{y}_i = f_i(y_j, u_k, t) \qquad j = 1, 2, \ldots, n \tag{1.1.3}$$

$$\psi_\ell (y_{io}, y_{if}, t_o, t_f) = 0 \qquad \ell = 1, 2, \ldots, p \leqq 2n + 2 \tag{1.1.4}$$

find the set which will minimize a sum of the form:

$$J' = g(y_{io}, y_{if}, t_o, t_f) + \int_{t_o}^{t_f} L(y_i(t), u_k(t), t)\, dt . \tag{1.1.5}$$

Here it is assumed that the functions f_j, L, g, and ψ_ℓ are of class C^2. In the above and throughout this presentation, a dot above a variable will be used to represent the derivative of the variable with respect to t, the independent variable. Likewise the subscripts o and f will indicate the evaluation of the variable or expression at the initial and final value of t, respectively. For the sake of brevity, the range of subscripts i, j, k, and ℓ will be as given above and will not be repeated in what follows.

Following the conventional method of Lagrange multipliers (Bryson and Ho, 1969), minimization of the augmented function

$$J^* = g + \mu_\ell \psi_\ell + \int_{t_o}^{t_f} [L - \lambda_i f_i + \lambda_i \dot{y}_i] \, dt \; . \tag{1.1.6}$$

is considered. In the above equation and throughout this report, repeated subscripts will be used to signify summation. Equation (1.1.6) was obtained by adjoining equations (1.1.3) and (1.1.4) to relation (1.1.5) as follows:

(a) multiplying relations (1.1.3) by the variables $\lambda_i(t)$, respectively, integrating from t_o to t_f and by adding the sum of the integrals to expression (1.1.5),

(b) multiplying equations (1.1.4) by the parameters μ_ℓ and adding the sum of the products to expression (1.1.5).

It is convenient to define the following functions:

$$G(y_{io}, y_{if}, t_o, t_f) = g + \mu_\ell \psi_\ell \tag{1.1.7}$$

$$H[y_i(t), \lambda_i(t), u_k(t), t] = \lambda_i f_i - L \; . \tag{1.1.8}$$

The function H is often referred to as the Hamiltonian. With these definitions, equation (1.1.6) may be written as

$$J^* = G + \int_{t_o}^{t_f} [-H + \lambda_i \dot{y}_i] \, dt \tag{1.1.9}$$

By considering small variations in the path and endpoints about a nominal path, it can be shown that if the functions $u_k(t)$ and $y_i(t)$

are a solution to the Problem of Bolza, then they must satisfy the following necessary conditions (Hestenes, 1966, pp. 346-351):

Condition I. There exist continuous multipliers $\lambda_i(t)$ and Hamiltonian function as defined by equation (1.1.8) such that:

(1) the Euler-Lagrange equations,

$$\dot{\lambda}_i = - \frac{\partial H}{\partial y_i} \tag{1.1.10}$$

$$\frac{\partial H}{\partial u_k} = 0 \tag{1.1.11}$$

are satisfied at every point along the path and,

(2) the transversality conditions

$$\frac{\partial G}{\partial t_o} + H_o = 0 \tag{1.1.12}$$

$$\frac{\partial G}{\partial y_{io}} - \lambda_{io} = 0 \tag{1.1.13}$$

$$\frac{\partial G}{\partial t_f} - H_f = 0 \tag{1.1.14}$$

$$\frac{\partial G}{\partial y_{if}} + \lambda_{if} = 0 \tag{1.1.15}$$

are satisfied by the endpoints.

Condition II. The inequality

$$H[y_i(t), \lambda_i(t), u_k^{no}(t), t] \leq H[y_i(t), \lambda_i(t), u_k(t), t] \tag{1.1.16}$$

must be satisfied for all $t_o \leq t \leq t_f$ and for all non-optimal control functions u^{no} (Weierstrass Condition).

Condition III. The k by k matrix

$$\frac{\partial^2 H}{\partial u_s \partial u_t} \qquad \begin{array}{l} s = 1, 2, \ldots, k \\ \vdots \\ t = 1, 2, \ldots, k \end{array} \tag{1.1.17}$$

must be negative semi-definite for a minimum (Legendre-Clebsch condition).

Condition IV. A fourth necessary condition is discussed by Bliss (1946, pp. 226-228) in classical dependent variable notation. He proves that the second order variation of a sum similar to J^* must be non-negative along a stationary arc, if that arc minimizes J^* Hestenes (1966, pp. 283-286) verifies this conclusion in modern control notation for the Problem of Bolza with fixed endpoints. As developed by Hestenes, the fourth necessary condition represents a necessary condition on the path alone; variations in the endpoints are not considered. If Condition III is satisfied, Condition IV is usually referred to as the Jacobi Condition. A further geometric interpretation of this condition is presented in the following section.

1.2 The Nature of Multiple Stationary Solutions

It is helpful in understanding the nature of multiple stationary solutions to classify them by the circumstances pertinent to their occurrence.

Fixed Endpoint Problems - Multiple stationary solutions are often obtained for problems with fixed endpoints, because only the first three necessary conditions have been applied. Bliss (1946, p. 235) has shown in dependent variable notation that the fourth necessary condition of Jacobi, taken together with the first three necessary conditions, suitably strengthened, forms a sufficient set of conditions for the Problem of Bolza.

Although analytically complex, the Jacobi condition has a simple geometric interpretation for problems with one state variable. In this case the set of all solutions forms a one dimensional family of extremal arcs, $y = y(t, c)$, which all pass through the initial point as shown in Figure (1.1). With each arc is associated a particular value of c. If arcs $y(t, c)$ and $y(t, c + \varepsilon)$ intersect in the limit as ε goes to 0, the point of intersection is called a conjugate point. The locus of such intersections is called the discriminant locus, also shown in Figure (1.1)

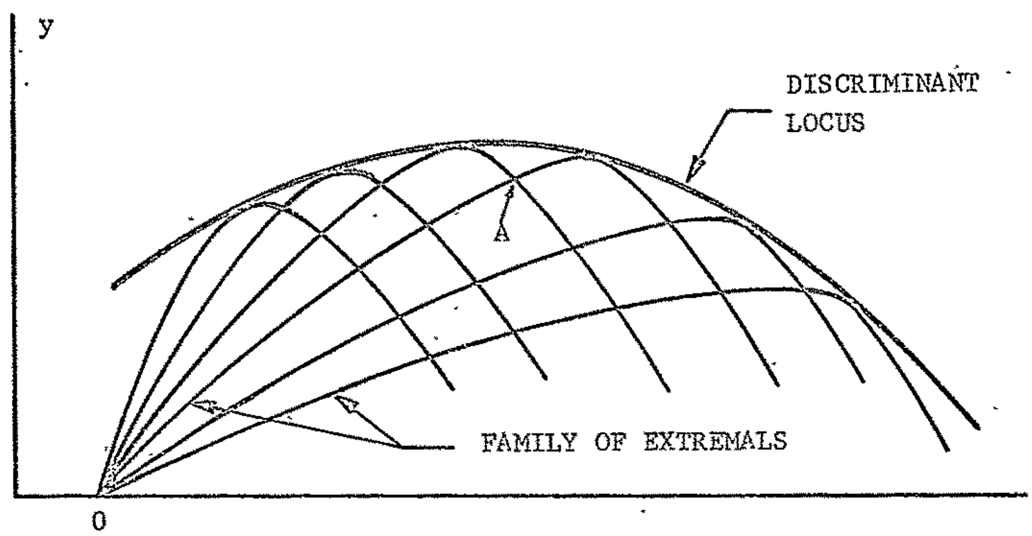

Fig. 1.1 A Family of Extremals and the Discriminant Locus

In terms of this geometry, the Jacobi condition requires that an optimal trajectory contain no conjugate point. Alternatively, the condition requires that an optimal solution may not touch the discriminant locus. Figure (1.1) shows that there are two solutions joining points O and A, one of which touches the discriminant locus and is therefore non-optimal. If the fourth necessary condition of Jacobi is applied in such cases of multiple stationary solutions, usually all but one of the trajectories will be shown to contain a point conjugate to the initial point, thus rendering them non-optimal.

Examples of this occurrence are many. The Brachistochrone problem with fixed endpoints graphically illustrates the idea. Consider the problem of a bead sliding down a wire under the influence of gravity alone. What should the shape of the wire be in order to minimize the time of transit between two points in a vertical plane? It is well known that the solution curves are cycloids. However, as shown in Figure (1.2), there are several different cycloids which satisfy the necessary conditions of the

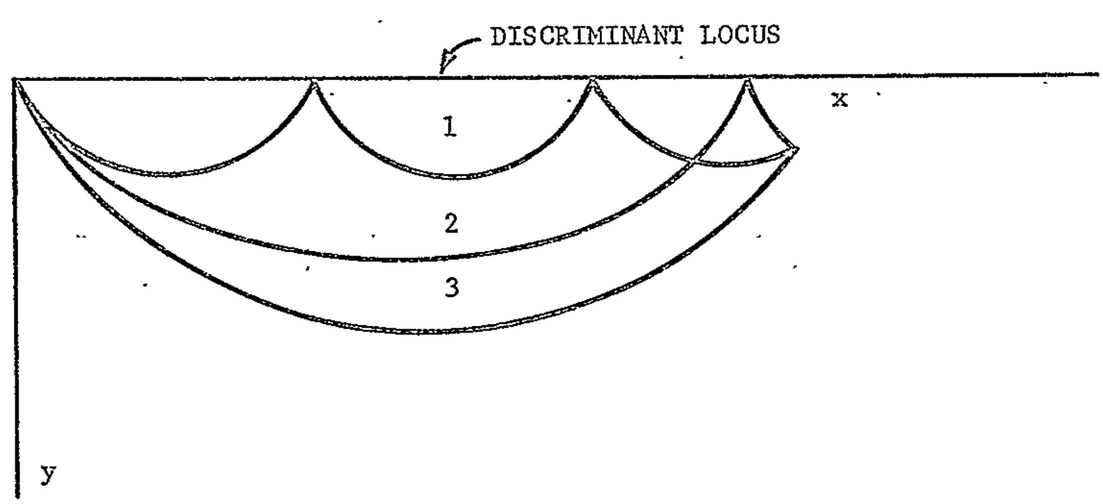

Fig. 1.2 Multiple Stationary Solutions for the Brachistochrone Problem

calculus of variations. It can be seen that the x-axis forms the discriminant locus and that the points where solutions 1 and 2 touch the discriminant locus are conjugate points. Since solutions 1 and 2 violate the Jacobi condition, it is evident that solution 3 is the true optimum.

In this case it has been possible to distinguish the true optimum from the candidates by applying a sufficiency condition pertaining to the field of extremal paths.

Variable Endpoint Problems - Problems with variable endpoints require that the endpoints of the trajectories, as well as the path, be selected in an optimum fashion. Consider the geodesics problem of trying to find the minimum distance from the origin to a given parabola as shown in Figure (1.3). It will be shown in Chapter 2 that two stationary solutions exist, viz., OA and OB. Once an endpoint has been selected, the problem becomes one of fixed endpoints,

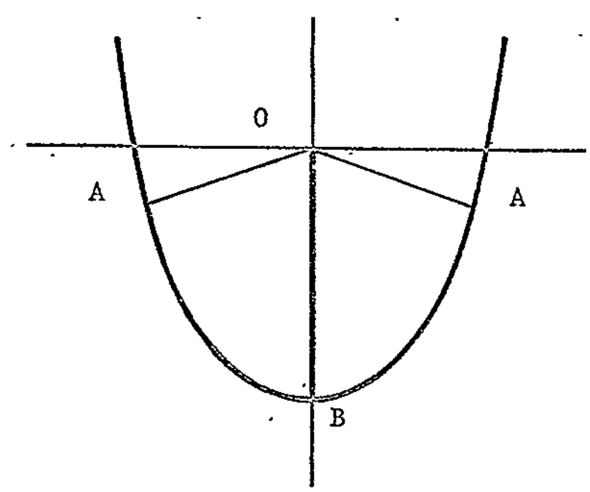

Fig. 1.3 Multiple Stationary Solutions to a Geodetics Problem

and the path sufficiency conditions previously discussed can then be applied. In this case, with the endpoint, A or B, thought of as being fixed, it can be shown that both solutions satisfy the Jacobi sufficiency condition (Bolza, 1961, pp. 84-86). The endpoints shown satisfy first order necessary conditions only, it is apparent that a second order endpoint condition may be useful in distinguishing the true optimal. A second order condition will be derived in Chapter 2. While the second order condition obtained is akin to the classical focal point condition (Bolza, 1961, pp. 104-10), the result is new in form and is directly applicable to the optimal control problem.

<u>Problems Requiring First and Second Order Conditions</u> - In the last two sections, the necessity of using fixed endpoint path sufficiency conditions and second order endpoint conditions was illustrated separately. It is not unusual, however, to encounter problems requiring the application of both the Jacobi condition and second order endpoint conditions to distinguish the true optimum from the set of multiple stationary solutions. To illustrate this situation, reconsider the Brachistrochrone problem where the final endpoint, instead of being fixed, is required to be on curve E as shown in Figure (1.4). Both trajectories OAB and OD satisfy second order endpoint conditions; that is, both solutions would represent a local minimum with respect to small variations of the endpoint along endpoint manifold E. As discussed before, point A is a conjugate point, thus violating the Jacobi condition. Trajectory OC violates a second order endpoint condition; it is in fact a local maximum with respect to small variations of the endpoint along E.

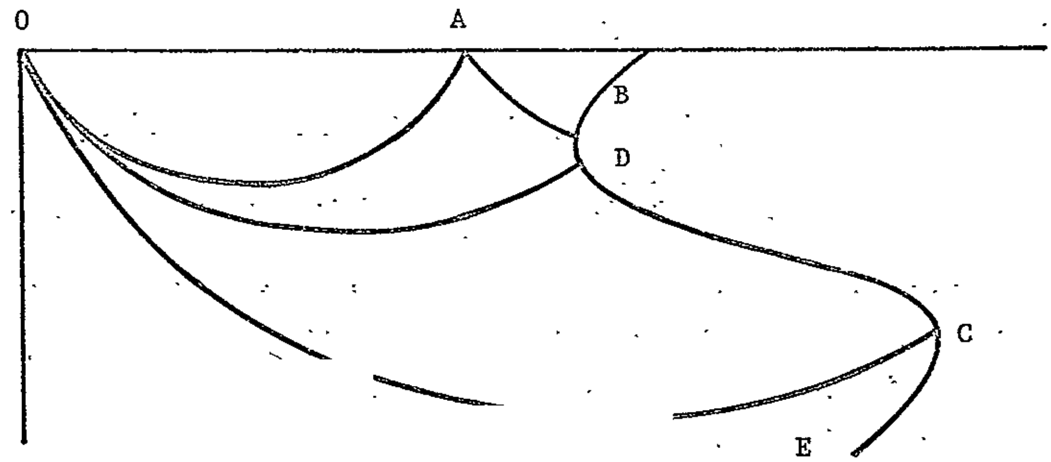

Fig. 1.4. A Problem Requiring Second Order Conditions

Thus through the use of both the Jacobi condition and a second order endpoint condition, OB is selected as the optimal candidate.

Problems with Periodic Solutions - Consider a system of equations (1.1.3) which exhibit periodic oscillations when no control effort is applied. It is not unusual for the optimal controls and adjoint variables of such a system to also demonstrate periodic motion with the same period, especially if the magnitude of the control is small.

The criterion by which the solution is terminated is also often periodic for problems exhibiting periodic oscillations. The terminating or cutoff condition is obtained from the transversality conditions (1.1.12) - (1.1.15) by eliminating the μ_ℓ parameters, to form a single relationship among the state and adjoint variables. The zeros of the cutoff function then represent the terminating condition. As an example, consider the problem of getting the mass of a thrusting harmonic oscillator to a specified height while minimizing the integral of the thrust with respect to time. A mass is connected in a parallel by a spring and dashpot to an inertial reference. The mass is capable of generating a bounded thrust in the upward direction. For simplicity it is assumed that the mass is constant. For small damping factors and null thrust, the state variables, position and velocity, the adjoint variables, and the cutoff function all exhibit damped periodic oscillations. As shown in Figure (1.5) the cutoff condition is satisfied during each period. For sufficiently small thrust amplitudes the cutoff function will deviate only slightly from that generated for null thrust, and will be satisfied at several points.

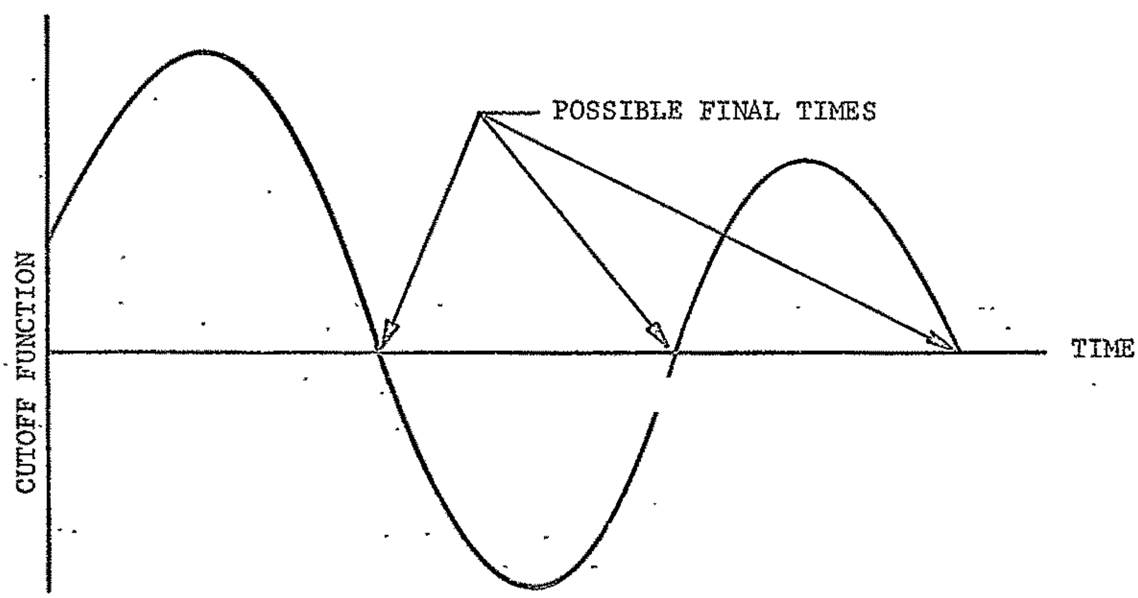

Fig. 1.5 Multiple Solutions Due to a Periodic Cutoff Function

Each time the cutoff condition is satisfied, a potential optimal endpoint and a corresponding stationary solution is obtained. Thus in periodic systems with weak bounded control, multiple stationary solutions may be encountered.

Problems with Singular Control - In cases where the Hamiltonian is linear in a bounded control variable, the control cannot be determined from the Euler-Lagrange equation (1.1.11). In this case the Hamiltonian can be written as

$$H = S(y_i, \lambda_i, t)u + Q(y_i, \lambda_i, t) \tag{1.2.1}$$

for scalar control. S is referred to as the switching function. The well-known Maximum Principle for problems with bounded control developed by Pontryagin et al. (1962) requires that

$$\begin{aligned} u &= u_{max} \quad \text{when } S > 0 \\ u &= u_{min} \quad \text{when } S < 0 \end{aligned} \tag{1.2.2}$$

However, for the case when $S \equiv 0$ over a non-vanishing time interval, the Maximum Principle is indeterminant, and u may take on intermediate values. This is the case of singular control. Leitmann (1966, pp. 57-58) has pointed out that, "While it is possible in a particular problem.... to rule out the possibility of [singular control], this cannot be done in general."

To demonstrate the existence of multiple stationary solutions in the case of singular control, examine the problem of minimizing

$$J = \frac{1}{2}\int_0^\infty x_1^2 \, dt \tag{1.2.3}$$

subject to the constraints:

$$\begin{aligned} \dot{x}_1 &= x_2 + u \qquad x_1(0) = x_{10} \qquad x_1(\infty) = 0 \\ \dot{x}_2 &= -u \qquad x_2(0) = x_{20} \qquad x_2(\infty) = 0 \end{aligned} \tag{1.2.4}$$

$$|u| \leq 1$$

This problem was first discussed by Johnson and Gibson (1963).

The Hamiltonian is

$$H = (\lambda_1 - \lambda_2)u + \lambda_1 x_2 - x_1^2/2 \tag{1.2.5}$$

In this case $S = \lambda_1 - \lambda_2$. If S is identically zero for a non-vanishing time interval, the possibility of singular control exists. By taking a suitable number of time derivatives, it can be shown that the singular control is given by $u = -x_1 - x_2$, and that the singular arcs are two lines $x_1(t) = 0$ and $s_1(t) + 2x_2(t) = 0$.

Figure (1.6) shows two possible stationary solutions to the problem starting at point A, one of which has a singular subarc. The first arc AB is the same for both solutions. In both solutions $u = -1$ along arc BC and then follow arc CO to the origin with $u = +1$. This is the so-called "bang-bang" solution. Alternately, at point B one may elect singular control, $u = x_2$, and proceed to the origin directly along arc BO.

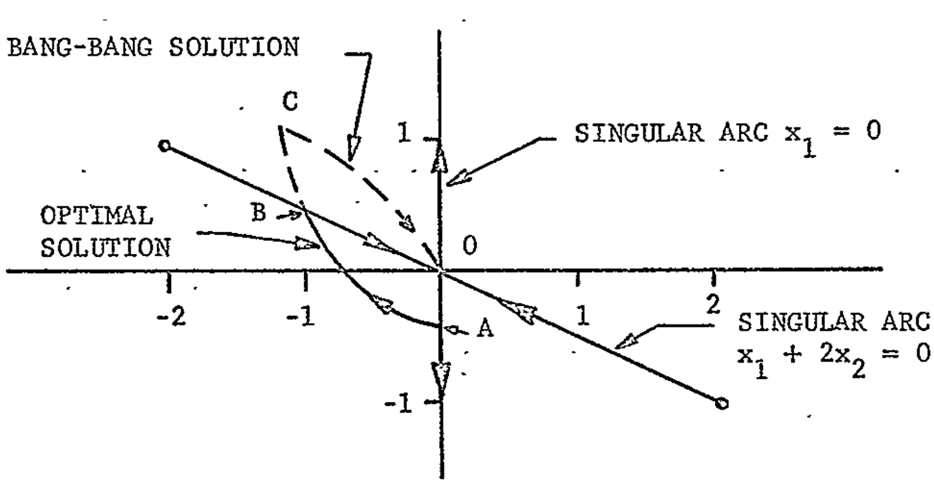

Fig. 1.6 Multiple Solutions Arising from Singular Control

Unfortunately, there is no guarantee that the solution with singular control is minimizing or that it will always enter the optimal solution, even if the possibility of singular solutions does exist. In this case the solution with bang-bang control, arc ABCO, has an index of performance almost 12 per cent larger than for the true optimal control which uses the singular control arc BO.

Recently, Kelly, Kopp and Moyer (1967) and Robbins (1965) have developed a new necessary condition for testing the optimality of singular subarcs.

A SECOND ORDER ENDPOINT CONDITION

In this section a second order condition is developed for variational problems with variable endpoints. Among the multiple stationary solutions that may exist, the condition provides a test for distinguishing those stationary solutions which are local minimums with respect to endpoint variations along the prescribed terminal manifold. The second order endpoint condition is related to sufficiency conditions of the classical calculus of variations in section 2.4. In section 2.5 the second order condition is illustrated with an example problem. Finally, a numerical algorithm is developed for applying the endpoint sufficiency condition to problems with no analytic solution.

2.1 Functional Relationships for the Problem of Bolza

Arguments used in Section 1 have implied that the conditions for the Problem of Bolza fall into two classes: those pertaining to the path and those pertaining to the endpoints. It has been further argued that conditions pertaining to the endpoints can be considered independently from those pertaining to path. Consider the Problem of Bolza as expressed in Section 1.1, equations (1.1.1) (1.1.5). In this formulation, and for the remainder of this section, the controls u_k are assumed to be unbounded functions of time. In addition, it is now assumed that the Jacobian is not equal to zero

$$\frac{\partial\left(\frac{\partial H}{\partial u_1}, \frac{\partial H}{\partial u_2}, \ldots, \frac{\partial H}{\partial u_k}\right)}{\partial\left(u_1, u_2, \ldots, u_k\right)} \neq 0 \tag{2.1.1}$$

for all points $(u_1, u_2, \ldots, u_k)$ in the control space. The function H has been previously defined in equation (1.1.8). If equation (2.1.1) is valid, the implicit function theorem (Buck, 1965, pp. 283-286) assures the existence of the k functional relations

$$u_k = u_k[y_i(t), \lambda_i(t)] \tag{2.1.2}$$

from the k control variable Euler-Lagrange equations (1.1.11). Condition (2.1.1) specifically eliminates from consideration those systems in which any state variable derivative, $\dot{y}$ as defined in equation (1.1.3), is a linear function of any of the control variables.

A solution to the Problem of Bolza is specified by solutions for the state variables y_i, as well as the control variables u_k, as functions of time. A selection of the initial time and the final time completes the solution. To obtain these solutions, the control variable Euler-Lagrange equations are first solved for the control variables u_k as functions of the state variables y_i and the adjoint variables λ_i. The functions f_i (1.1.3) and the L function (1.1.5) are now explicitly dependent only on the state variables, the adjoint variables, and time.

$$f_i = f_i[y_j(t);\ u_k(y_j(t), \lambda_j(t));\ t] \tag{2.1.3}$$

$$L = L[y_j(t);\ u_k(y_j(t),\ \lambda_j(t));\ t] \tag{2.1.4}$$

Similarly the H function becomes an explicit function of the state variables, the adjoint variables, the adjoint variables, and time, alone:

$$H^o = H^o[y_i(t);\ \lambda_i(t);\ u_k(y_i(t),\ \lambda_i(t));\ t] \tag{2.1.5}$$

Finally, it is evident that a similar functional relationship exists for the time derivatives of adjoint variables,

$$\dot{\lambda}_i = P_i[y_i(t),\ \lambda_i(t),\ t] = -\frac{\partial H}{\partial y_i} \tag{2.1.6}$$

In summary, once the optimal control is selected, the state variable differential equations (1.1.3) and the adjoint variable differential equations (1.1.10) comprise a set of 2n first order nonlinear differential equations in the 2n state and adjoint variables and time. This set of differential equations can be integrated in theory, yielding

$$y_i = y_i(t,\ c_r) \qquad r = 1,\ 2,\ \ldots,\ 2n \tag{2.1.7}$$

$$\lambda_i = \lambda_i(t,\ c_r) \tag{2.1.8}$$

where the c_r's are constants of integration. The p end conditions (1.1.4) and the 2n + 2 equations representing the transversality necessary conditions (1.1.12) - (1.1.15) comprise a set of (2n + p + 2) non-linear algebraic equations in the 2n constants c_r, the p parameters μ_ℓ, the initial time t_o, and the final time t_f.

The initial values (y_{io}, λ_{if}, t_f) are specified. Hence, the c_r's may be determined as a function of these initial and/or final values by evaluating equations (2.1.7) and (2.1.8) at either the initial or final point. For example, a solution for the c_r's would be specified by the set $(y_{io}, \lambda_{io}, t_o, t_f)$, the set $(y_{if}, \lambda_{if}, t_f, t_o)$ or the set $(y_{io}, y_{if}, t_o, t_f)$. While it is difficult to attach any physical meaning to the initial or final values of the Lagrange multipliers, the initial and final values of the state variables have an immediate physical significance. For this reason, the state variable endpoints have been selected to functionally represent the c_r constants of integration for the rest of this section. Thus equations (.1.7) and (2.1.8) will be written as

$$y_j = y_j(t, y_{io}, y_{if}, t_o, t_f) \qquad (2.1.9)$$

and $$t_o < t < t_f$$

$$\lambda_j = \lambda_j(t, y_{io}, y_{if}, t_o, t_f) \qquad (2.1.10)$$

By substituting the functional relationships exhibited in equations (2.1.9) and (2.1.10) into relations (2.1.2) - (2.1.6), it can be seen that the functions u_k, L, f_i, H^o, and P_i can all be written as explicit functions of the set $(t, y_{io}, y_{if}, t_o, t_f)$. These functional relations, together with that for the function G from equation (1.1.7) are summarized for reference below:

$$u_k = u_k\,(t, y_{io}, y_{if}, t_o, t_f) \qquad (2.1.11)$$

$$L = L\,(t, y_{io}, y_{if}, t_o, t_f) \qquad (2.1.12)$$

$$\dot{y}_j = f_j(t, y_{io}, y_{if}, t_o, t_f) \qquad (2.1.13)$$

$$\lambda_j = P_j(t, y_{io}, y_{if}, t_o, t_f) \qquad (2.1.14)$$

$$H^o = H^o[y_j(t, y_{io}, y_{if}, t_o, t_f); \lambda_j(t, y_{io}, y_{if}, t_o, t_f); u_k(t, y_{io}, y_{if}, t_o, t_f); t] \tag{2.1.15}$$

$$G = G(\mu_\ell, y_{io}, y_{if}, t_o, t_f) \tag{2.1.16}$$

So that there will be no confusion as to the meaning of the subscripts, note that

$$y_{jo} \equiv y_j \Big|_{t=t_o} \tag{2.1.17}$$

$$\lambda_{jo} \equiv \lambda_j \Big|_{t=t_o} \tag{2.1.18}$$

$$y_{jf} \equiv y_j \Big|_{t=t_f} \tag{2.1.19}$$

$$\lambda_{jf} \equiv \lambda_j \Big|_{t=t_f} \tag{2.1.20}$$

Using the functional relationships summarized above form the augmented function $J^* = J + \mu_\ell \psi_\ell$ where

$$J(y_{io}, y_{if}, t_o, t_f) = g(y_{io}, y_{if}, t_o, t_f) \tag{2.1.21}$$

$$+ \int_{t_o}^{t_f} \Big[-H^o(t, y_{io}, y_{if}, t_o, t_f)$$

$$+ \lambda_j(t, y_{io}, y_{if}, t_o, t_f)\, \dot{y}_j(t, y_{io}, y_{if}, t_o, t_f)\Big]\, dt$$

By requiring the trajectory to satisfy certain necessary conditions regarding path, equations (1.1.10) and (1.1.11), the Problem of Bolza has been reduced to the problem of minimizing J, a function of endpoints, subject to the ψ_ℓ algebraic constraints on the endpoints.

Before proceeding with the minimization of J, it is appropriate to consider a graphical interpretation of the functional relationship for the state variables expressed in equation (2.1.9). Figure 2.1 shows a general state function $y_i(t, y_{io}, y_{if}, t_o, t_f)$ as a function of time. From the figure it can be seen that a change in the final state Δy_{if} while holding all of the other endpoints fixed causes a change in y_i for all values of t. Likewise a change in the final time Δt_f while holding all of the other endpoints fixed causes a change in the state y_i for all values of t.

Formalizing this graphical interpretation in terms of differentials yields results which will be of value in the following sections. Using equation (2.1.9), the differential of the state variables may be written as,

$$dy_i(t, y_{jo}, y_{jf}, t_o, t_f) = \frac{\partial y_i}{\partial t} dt + \frac{\partial y_i}{\partial t_o} dt_o + \frac{\partial y_i}{\partial t_f} dt_f + \frac{\partial y_i}{\partial y_{jo}} dy_{jo} + \frac{\partial y_i}{\partial y_{jf}} dy_{jf} \tag{2.1.22}$$

Evaluating this expression at $t = t_f$ gives

$$dy_i(t_f, y_{jo}, y_{jf}, t_o, t_f) = dy_{if} = \frac{\partial y_i}{\partial t}\bigg|_f dt_f + \frac{\partial y_i}{\partial t_o}\bigg|_f dt_o + \frac{\partial y_i}{\partial t_f}\bigg|_f dt_f + \frac{\partial y_i}{\partial y_{jo}}\bigg|_f dy_{jo} + \frac{\partial y_i}{\partial y_{jf}}\bigg|_f dy_{jf} \tag{2.1.23}$$

The sum represented by the last term in the above equation can be separated into those products for which $i \neq j$ and that for which $i = j$. Transposing dy_{if} to the right hand side, equation (2.1.23)

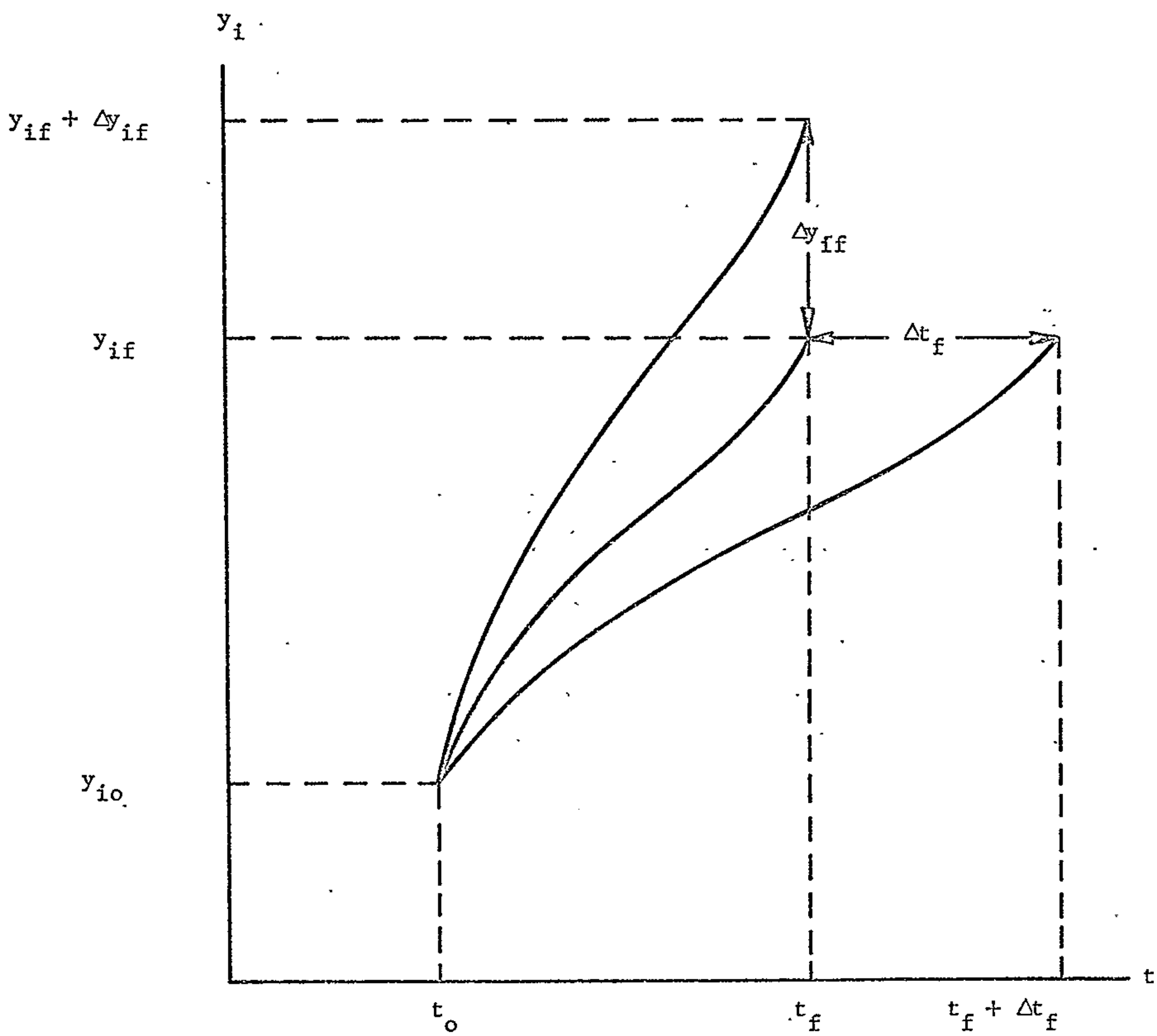

Fig. 2.1 A Representation of y_i as a Function of y_{if} and t_f,

becomes

$$0 = \left[\frac{\partial y_i}{\partial t}\bigg|_f + \frac{\partial y_i}{\partial t_f}\bigg|_f \right] dt_f + \frac{\partial y_i}{\partial y_{jo}}\bigg|_f dy_{jo}$$

$$+ \frac{\partial y_i}{\partial y_{jf}}\bigg|_{\substack{f \\ i \neq j}} dy_{if} + \left[\frac{\partial y_i}{\partial y_{jf}}\bigg|_f \quad 1 \right] dy_{\substack{jf \\ i=j}} \tag{2.1.24}$$

In the above equation the repeated subscripts on the last term do not imply summation. Since t, t_o, t_f, y_{io}, and y_{if} have been assumed to be independent, equation (2.1.24) implies that

$$\frac{\partial y_i}{\partial t}\bigg|_f = -\frac{\partial y_i}{\partial t_f}\bigg|_f \tag{2.1.25}$$

$$\frac{\partial y_i}{\partial t_o}\bigg|_f = 0 \tag{2.1.26}$$

$$\frac{\partial y_i}{\partial y_{jo}}\bigg|_f = 0 \tag{2.1.27}$$

$$\frac{\partial y_i}{\partial y_{jf}}\bigg|_{\substack{f \\ i \neq j}} = 0 \tag{2.1.28}$$

$$\frac{\partial y_i}{\partial y_{jf}}\bigg|_{\substack{f \\ i = j}} = 1 \tag{2.1.29}$$

By evaluating equation (2.1.22) at $t = t_o$ and following arguments similar to the ones above, it can be shown that

$$\left.\frac{\partial y_i}{\partial t}\right|_o = -\left.\frac{\partial y_i}{\partial t_f}\right|_o \tag{2.1.30}$$

$$\left.\frac{\partial y_i}{\partial t_f}\right|_o = 0 \tag{2.1.31}$$

$$\left.\frac{\partial y_i}{\partial y_{jf}}\right|_o = 0 \tag{2.1.32}$$

$$\left.\frac{\partial y_i}{\partial y_{jo}}\right|_{\substack{o \\ i \neq j}} = 0 \tag{2.1.33}$$

$$\left.\frac{\partial y_i}{\partial y_{jo}}\right|_{\substack{o \\ i=j}} = 1 \tag{2.1.34}$$

These identities will be useful in the proofs of necessary and sufficient conditions in the next section.

2.2 Derivation of Transversality Conditions

In determining the functional relationships in the last section, it was assumed that the control and adjoint variables were chosen so as to satisfy the Euler-Lagrange Equations (1.1.10) and (1.1.11). Equations (1.1.10) and (1.1.11) are referred to as the first path necessary conditions. In this section the endpoint necessary conditions (transversality conditions) are derived assuming that the first necessary conditions for path are satisfied.

The solution to the path necessary conditions determines one or more trajectories (see section 1.3), any of which may be expressed functionally as a set $[y_j(t, y_{io}, y_{if}, t_o, t_f), \lambda_j(t, y_{io}, y_{if}, t_o, t_f),$ $u_k(t, y_{io}, y_{if}, t_o, t_f)]$ as shown in section 2.1. Once the functions representing one of these trajectories is substituted into the integral in equation (2.1.21), the integration can be performed. It is therefore clear that once the trajectory is specified, J is a function of only the parameters y_{io}, y_{if}, t_o and t_f. Specifying the path reduces the problem of minimizing J to the well-known problem of finding the minimum of a function of several variables subject to algebraic equations of constraint (Bryson and Ho, 1969).

It is shown in Appendix A that if the arguments of J in equation (2.1.21) are to satisfy the constraints and minimize J, then it is necessary that the partial derivatives of the auxiliary function, shown below, with respect to y_{io}, y_{if}, t_o, and t_f all be equal to zero.

The J function is defined by $J^* = J + \mu_\ell \psi_\ell$ where J is given by equation (2.1.21). Using the definition of the function G from equation (1.2.8), J^* may be functionally represented as

$$J^*[y_{io}, y_{if}, t_o, t_f, \mu_\ell] = G[y_{io}, y_{if}, t_o, t_f, \mu_\ell] + \int_{t_o}^{t_f} \left[-H^o(y_i, \lambda_i, u_k^o, t) + \lambda_i \frac{\partial y_i}{\partial t}\right] dt \tag{2.2.1}$$

In the above equation it is understood that y_i, λ_i, and u_k are all functions of the set $(t, y_{io}, y_{if}, t_o, t_f)$. In writing the functional relationship shown above, it has been assumed that the controls u_k have been chosen in an optimal fashion in accordance with the control variable

Euler-Lagrange Equation (1.1.11). This is indicated by the superscript o on u_k and on H. The partial derivative of J^* with respec' to y_{io} can now be written:

$$\frac{\partial J^*}{\partial y_{jo}} = \frac{\partial G}{\partial y_{jo}} + \int_{t_o}^{t_f} \left[- \frac{\partial H}{\partial y_{jo}} + \frac{\partial \lambda_i}{\partial y_{jo}} \frac{\partial y_i}{\partial t} + \lambda_i \frac{\partial^2 y_i}{\partial y_{jo} \partial t} \right] dt \tag{2.2.2}$$

Here Leibnitz Rule (Hildebrand, 1948, p. 360) has been used for differentiation of an integral with respect to a parameter. Using the identity

$$\frac{d}{dt} \left[\lambda_i \frac{\partial y_i}{\partial y_{jo}} \right] = \lambda_i \frac{\partial^2 y_i}{\partial y_{jo} \partial t} + \frac{\partial \lambda_i}{\partial t} \frac{\partial y_i}{\partial y_{jo}} \tag{2.2.3}$$

and expanding $\frac{\partial H}{\partial y_{jo}}$, equation (2.2.2) may be written as

$$\begin{aligned} \frac{\partial J^*}{\partial y_{jo}} = \frac{\partial G}{\partial y_{jo}} + \int_{t_o}^{t_f} \Bigg[&- \frac{\partial H}{\partial y_i} \frac{\partial y_i}{\partial y_{jo}} - \frac{\partial H}{\partial \lambda_i} \frac{\partial \lambda_i}{\partial y_{jo}} - \frac{\partial H}{\partial u_k} \frac{\partial u_k}{\partial y_{jo}} \\ &+ \frac{\partial \lambda_i}{\partial y_{jo}} \frac{\partial y_i}{\partial t} - \frac{\partial \lambda_i}{\partial t} \frac{\partial y_i}{\partial y_{jo}} \Bigg] dt + \left[\lambda_i \frac{\partial y_i}{\partial y_{jo}} \right]_{t_o}^{t_f} \end{aligned} \tag{2.2.4}$$

Terms under the integral sign may be combined to give

$$\begin{aligned} \frac{\partial J^*}{\partial y_{jo}} = \frac{\partial G}{\partial y_{jo}} &+ \lambda_i \Big|_{t_f} \frac{\partial y_i}{\partial y_{jo}} \Big|_{t_f} - \lambda_i \Big|_{t_o} \frac{\partial y_i}{\partial y_{jo}} \Big|_{t_o} \\ &+ \int_{t_o}^{t_f} \Bigg[- \left(\frac{\partial H}{\partial \lambda_i} - \frac{\partial y_i}{\partial t} \right) \frac{\partial \lambda_i}{\partial y_{jo}} - \left(\frac{\partial H}{\partial y_i} + \frac{\partial \lambda_i}{\partial t} \right) \frac{\partial y_i}{\partial y_{jo}} \\ &+ \frac{\partial H}{\partial u_k} \frac{\partial u_k}{\partial y_{jo}} \Bigg] dt \end{aligned} \tag{2.2.5}$$

Note from equations (2.1.9) and (2.1.10) that once the optimal endpoints have been selected,

$$\frac{\partial y_i}{\partial t} = \frac{dy_i}{dt} \quad \text{and} \quad \frac{\partial \lambda_i}{\partial t} = \frac{d\lambda_i}{dt}$$

The integral term vanishes, since equations (1.1.3), (1.1.10) and (1.1.11) were used to generate the functional relations (2.1.9) and (2.1.10).

Using equations (2.1.27), (2.1.33) and (2.1.34), it can be concluded that the sums represented by the two remaining terms not containing G in equation (2.2.5) reduce to a single term, $-\lambda_j\Big|_{t_o}$. With these considerations, equation (2.2.5) reduces to

$$\frac{\partial J^*}{\partial y_{jo}} = \frac{\partial G}{\partial y_{jo}} - \lambda_j\Big|_{t_o} = 0 \tag{2.2.6}$$

By taking the derivative of J^* with respect to y_{jf} and using arguments similar to those just presented (in this case equations (2.1.32), (2.1.28) and (2.1.29) must be taken into account), it can be shown that

$$\frac{\partial J^*}{\partial y_{jf}} = \frac{\partial G}{\partial y_{jf}} + \lambda_j\Big|_{t_f} = 0 \tag{2.2.7}$$

Two more necessary conditions remain to be derived. These result from taking the partial derivatives of J^* with respect to the remaining two variables, t_o and t_f. Performing the first of these operations yields

$$\frac{\partial J^*}{\partial t_o} = \frac{\partial G}{\partial t_o} + \int_{t_o}^{t_f} \left[- \frac{\partial H}{\partial t_o} + \frac{\partial \lambda_i}{\partial t_o} \frac{\partial y_i}{\partial t} + \lambda_i \frac{\partial y_i}{\partial t \partial t_o} \right] dt$$

$$- \left[- H + \lambda_i \frac{\partial y_i}{\partial t} \right]_{t_o} \tag{2.2.8}$$

Here again Leibnitz Rule has been used; this time the limits of integration are functions of the differentiating variable. Using the identity

$$\frac{d}{dt} \left[\lambda_i \frac{\partial y_i}{\partial t_o} \right] = \lambda_i \frac{\partial y_i}{\partial t \partial t_o} + \frac{\partial \lambda_i}{\partial t} \frac{\partial y_i}{\partial t_o} \tag{2.2.9}$$

and expanding $\frac{\partial H}{\partial t_o}$, equation (2.2.8) may be written as

$$\frac{\partial J^*}{\partial t_o} = \frac{\partial G}{\partial t_o} + \int_{t_o}^{t_f} \left[- \frac{\partial H}{\partial y_i} \frac{\partial y_i}{\partial t_o} - \frac{\partial H}{\partial \lambda_i} \frac{\partial \lambda_i}{\partial t_o} - \frac{\partial H}{\partial u_k} \frac{\partial u_k}{\partial t_o} \right.$$

$$\left. + \frac{\partial \lambda_i}{\partial t_o} \frac{\partial y_i}{\partial t} - \frac{\partial \lambda_i}{\partial t} \frac{\partial y_i}{\partial t_o} \right] dt + \left[\lambda_i \frac{\partial y_i}{\partial t_o} \right]_{t_o}^{t_f} \tag{2.2.10}$$

$$- \left[- H + \lambda_i \frac{\partial y_i}{\partial t} \right] \Bigg|_{t_o}$$

Terms outside the integral may be evaluated at the endpoints indicated and terms under the integral sign combined to give

$$\frac{\partial J^*}{\partial t_o} = \frac{\partial G}{\partial t_o} + H \Big|_{t_o} - \lambda_i \Big|_{t_o} \frac{\partial y_i}{\partial t} \Big|_{t_o} + \lambda_i \Big|_{t_f} \frac{\partial y_i}{\partial t_o} \Big|_{t_f} - \lambda_i \Big|_{t_o} \frac{\partial y_i}{\partial t_o} \Big|_{t_o}$$

$$+ \int_{t_o}^{t_f} \left[- \left(\frac{\partial H}{\partial y_i} + \frac{\partial \lambda_i}{\partial t} \right) \frac{\partial y_i}{\partial t_o} - \left(\frac{\partial H}{\partial \lambda_i} - \frac{\partial y_i}{\partial t} \right) \frac{\partial \lambda_i}{\partial t_o} - \left(\frac{\partial H}{\partial u_k} \right) \frac{\partial u_k}{\partial t_o} \right] dt \tag{2.2.11}$$

The integral again vanishes identically for optimal paths. Using equations (2.1.26) and (2.1.30), the three terms outside the integral representing summations can also be equated to zero. With these observations, equation (2.2.11) reduces to

$$\frac{\partial J^*}{\partial t_o} = \frac{\partial G}{\partial t_o} + H \Big|_{t_o} = 0 \tag{2.2.12}$$

By taking the derivative of J^* with respect to t_f, following a line of reasoning similar to that just given, and using equations (2.1.31) and (2.1.25), it can be shown that

$$\frac{\partial J^*}{\partial t_f} = \frac{\partial G}{\partial t_f} - H \Big|_{t_f} = 0 \tag{2.2.13}$$

These results are summarized in the following statement:

Transversality Necessary Condition for Endpoints - If a trajectory satisfies the Euler-Lagrange and state variable differential equations, equations (1.1.10), (1.1.11), and (1.1.3), and if the set $E = [y_{io}, y_{if}, t_o, t_f, \mu_\ell]$ satisfies endpoint equations of constraint (1.1.4) and provides a local minimum of J with respect to small allowable variations in the endpoints, then the set E must satisfy equations (2.2.6), (2.2.7), (2.2.12) and (2.2.13).

These latter equations are referred to as the endpoint necessary conditions or, classically, as the transversality necessary conditions.

2.3 Derivation of Second Order Endpoint Conditions

In the last section the function J was shown to be a function of the endpoint variables y_{io}, y_{if}, t_o, and t_f when evaluated along an optimal path. The function J is constrained, however, through the p equations of constraint ψ_ℓ of equation (1.1.4). Second order conditions for determining the minimum of a function whose arguments must satisfy algebraic equations of constraint has recently been discussed by Vincent and Cliff (1970, pp. 171-173). Their methods will be used here. For reference, a detailed discussion of the minimization of a function of several variables is included in Appendix A.

Before presenting a statement of the second order condition, a brief discussion and definition of notation are in order. Since the algebraic equations of constraint for the Problem of Bolza define relationships among the endpoint variables, the endpoint variables are not all independent. Since there are p equations of constraint and (2n + 2) endpoint variables, there are only (2n - p + 2) independent endpoint variables. The p dependent variables are determined by the p equations of constraint. Any p of the variables can be considered to be the dependent variables. The choice is one of convenience Let the p dependent variables be denoted by the column vector $\underline{w}$ and the remaining (2n - p + 2) independent variables be denoted by the column vector $\underline{v}$. Let the vector $\underline{\psi}$ represent a vector whose elements are the ψ_ℓ constraint functions. Equation (2.3.1) summarizes these relations.

$$\underline{\psi} = \begin{vmatrix} \psi_1(\underline{w}, \underline{v}) \\ \psi_2(\underline{w}, \underline{v}) \\ \vdots \\ \psi_3(\underline{w}, \underline{v}) \end{vmatrix} \qquad \underline{w} = \begin{vmatrix} w_1 \\ w_2 \\ \vdots \\ w_p \end{vmatrix} \qquad v = \begin{vmatrix} v_1 \\ v_2 \\ \vdots \\ v_q \end{vmatrix} \tag{2.3.1}$$

$$q = 2n - p + 2$$

The identification of the elements of $\underline{v}$ and the elements of $\underline{w}$ with the endpoints y_{io}, y_{if}, t_o, and t_f is arbitrary except that the set of ψ_ℓ equations must contain every element of $\underline{w}$ and, in addition, every ψ_ℓ equation must contain at least one element of $\underline{w}$. It is convenient to define an additional column vector $\underline{r}$, whose first elements are the dependent variables and last elements the independent variables:

$$\underline{r} = \begin{vmatrix} w_1 \\ \cdot \\ \cdot \\ \cdot \\ w_p \\ v_1 \\ \cdot \\ \cdot \\ \cdot \\ v_q \end{vmatrix} \tag{2.3.2}$$

With these vectors define $\left[\frac{\partial J^*}{\partial \underline{r} \partial \underline{r}} \right]$ as a (2n + 2) by (2n + 2) matrix with elements $a_{ij} = \frac{\partial J^*}{\partial r_i \partial r_j}$. Let the matrix Φ be defined by

$$\Phi = - \left[\frac{\partial \underline{\psi}}{\partial \underline{w}} \right]^{-1} \left[\frac{\partial \underline{\psi}}{\partial \underline{v}} \right] \tag{2.3.3}$$

where $\left[\frac{\partial \underline{\psi}}{\partial \underline{w}} \right]$ is a p by p matrix with elements $a_{ij} = \frac{\partial \psi_i}{\partial w_j}$ and $\left[\frac{\partial \underline{\psi}}{\partial \underline{v}} \right]$ is a p by q matrix with elements $a_{ij} = \frac{\partial \psi_i}{\partial v_j}$. It is shown in Appendix A that the Φ matrix is the linear transformation which transforms differential changes in the independent variables into differential changes in the dependent variables. The Φ matrix has p rows and q columns. Finally, define the (2n + 2) by q partitioned matrix Ω as

$$\Omega = \left[\frac{\Phi}{I}\right] \tag{2.3.4}$$

where I represents a q by q identity matrix.

With these definitions the second order endpoint condition may now be stated:

Second Order Endpoint Condition. If E represents a set of endpoints and multipliers $[y_{io}, y_{if}, t_o, t_f, \mu_\ell]$ which satisfy the transversality necessary condition for endpoints, and if the set E represents a local interior minimum of the function J with respect to small allowable variations in the endpoints then the quadratic form

$$\underline{dv}^T \Omega^T \left[\frac{\partial J^*}{\partial \underline{r} \partial \underline{r}} \right] \Omega \quad \underline{dv} \tag{2.3.5}$$ †

in the differentials $\underline{dv}$ is positive semi-definite when evaluated at the stationary point E.

To implement this test, it is necessary to evaluate the elements of the matrix Φ and the elements of the matrix $\left[\frac{\partial J^*}{\partial \underline{r} \partial \underline{r}} \right]$. Evaluation of elements of Φ represents no problem since the functional form of the constraints is specified in the problem statement. However, the analytic evaluation of the second partial derivatives of J^* with respect to the endpoints is not so simple.

The second partial derivatives of J^* can be obtained by taking the

---

† Conversely if the transversality conditions hold and if (2.3.5) is positive definite then the set E is a local interior minimum (sufficiency).

partial derivatives of the transversality necessary conditions with respect to the endpoints r_i.

$$\frac{\partial}{\partial r_i}\left(\frac{\partial J^*}{\partial y_{jo}}\right) = \frac{\partial G}{\partial r_i \partial y_{jo}} - \frac{\partial}{\partial r_i}\left(\lambda_j\Big|_{t_o}\right) \tag{2.3.6}$$

$$\frac{\partial}{\partial r_i}\left(\frac{\partial J^*}{\partial y_{jf}}\right) = \frac{\partial G}{\partial r_i \partial y_{jf}} + \frac{\partial}{\partial r_i}\left(\lambda_j\Big|_{t_f}\right) \tag{2.3.7}$$

$$\frac{\partial}{\partial r_i}\left(\frac{\partial J^*}{\partial t_o}\right) = \frac{\partial G}{\partial r_i \partial t_o} + \frac{\partial}{\partial r_i}\left(H\Big|_{t_o}\right) \tag{2.3.8}$$

$$\frac{\partial}{\partial r_i}\left(\frac{\partial J^*}{\partial t_f}\right) = \frac{\partial G}{\partial r_i \partial t_f} - \frac{\partial}{\partial r_i}\left(H\Big|_{t_f}\right) \tag{2.3.9}$$

where in the above equations $i = 1, 2, \ldots, 2n + 2$. The functional form of G as a function of the endpoints is specified by the statement of the problem. However, the functions λ_j and H are not known functions of the endpoints until the state variable and Euler-Lagrange differential equations have been integrated analytically.

Since analytical integration is often difficult or impossible, it would be desirable to evaluate the partial derivatives of λ_i and H with respect to the endpoints in terms of functional forms specified in the statement of the problem. A complete set of relationships of this type were not found. Unless future investigation establishes such relationships, analytic application of the second order condition requires an analytic solution of the state variable and Euler-Lagrange differential equations.

Some interesting relations of this type are easily obtained however. Each of the elements of the matrix $\left[\frac{\partial J^*}{\partial \underline{r} \partial \underline{r}}\right]$ is composed of a sum of a second term. The matrix can therefore be expressed as the sum of two matrices,

$$\left[\frac{\partial J^*}{\partial \underline{r} \partial \underline{r}} \right] = \left[\frac{\partial G}{\partial \underline{r} \partial \underline{r}} \right] + A \tag{2.3.10}$$

where the matrix A is determined from equations (2.3.6) - (2.3.9) Since J^* and G are of class C^2 by hypothesis, both J^* and G must be symmetric about their major diagonals. The obvious conclusion is that matrix A must also by symmetric. By equating symmetric elements of A, the following identities can be established:

$$\frac{\partial \lambda_{io}}{\partial y_{jo}} = \frac{\partial \lambda_{jo}}{\partial y_{io}} \qquad \frac{\partial \lambda_{if}}{\partial y_{jf}} = \frac{\partial \lambda_{jf}}{\partial y_{if}} \tag{2.3.11}$$

$$\frac{\partial H_o}{\partial y_{io}} = \frac{\partial \lambda_{io}}{\partial t_o} \qquad \frac{\partial H_f}{\partial y_{if}} = -\frac{\partial \lambda_{if}}{\partial t_f} \tag{2.3.12}$$

$$\frac{\partial H_f}{\partial y_{io}} = \frac{\partial \lambda_{io}}{\partial t_f} \qquad \frac{\partial H_o}{\partial y_{if}} = \frac{\partial \lambda_{if}}{\partial t_o} \tag{2.3.13}$$

$$\frac{\partial \lambda_{if}}{\partial y_{jo}} = \frac{\partial \lambda_{jo}}{\partial y_{if}} \qquad \frac{\partial H_f}{\partial t_o} = -\frac{\partial H_o}{\partial t_f} \tag{2.3.14}$$

In addition, the following relations can be established by considering the functional relationships exhibited in section 2.1.

$$\frac{\partial H_f}{\partial t_f} = \frac{\partial H}{\partial t} + f_{if} \frac{\partial \lambda_{if}}{\partial t_f} \tag{2.3.15}$$

$$\frac{\partial H_f}{\partial y_{if}} = \frac{\partial \lambda_{if}}{\partial t_f} + f_{jf} \frac{\partial \lambda_{jf}}{\partial y_{if}} \tag{2.3.16}$$

Similar equations exist for the initial point.

Unfortunately, a sufficient number of these relationships have not been found to determine the elements of A in terms of known functions in the problem statement. The determination of further relationships and the ultimate determination of the elements of A without resort to analytical integration of the state variable and Euler-Lagrange equations poses an interesting problem for future investigations.

2.4 Relation to Classical Theory

Bolza (1961, pp. 102-103) gives an excellent summary of the various classical approaches to the development of necessary and sufficient conditions for variable endpoint problems. Because of the pertinence of his remarks to this presentation, his historical synopsis is quoted in detail:

> Three essentially different methods have been proposed for the discussion of problems with variable endpoints:
>
> 1. The method of the Calculus of Variations proper: † It consists in computing δJ and $\delta^2 J$ either by means of Taylor's formula or by the method of differentiation with respect to , ...and discussing the conditions $\delta J=0$ $\delta^2 J \geqq 0$ ...
>
> 2. The method of Differential Calculus: This method is explained in general way in Dienger's Grundriss des Variationsrechnung (1867). It decomposes the problem into two problems by first considering variations which leave the endpoints fixed, and then variations which vary the endpoints, the neighboring curves considered being themselves extremals. The second part of the problem reduces to a problem of the theory of ordinary maxima and minima. This method has been used by A. Mayer in an earlier paper on the second variations in the case of variable endpoints for the general type of integrals mentioned above (Leipziger Berichte (1884), page 99).

---

† The first and second order variation of the integral are written as δJ and $\delta^2 J$, respectively. Variations in the endpoints and in the path are considered simultaneously in this method.

It is superior to the first method not only on account of its greater simplicity and its more elementary character, but because by utilizing the well-known sufficient conditions for ordinary maxima and minima it leads, in a certain sense, to sufficient conditions it combined with Weierstrauss's sufficient conditions for the case of fixed endpoints....

3. Kneser's method: This method, which has been developed by Kneser in his Lehrbuch†, is based upon an extension of certain well-known theorems on geodesics. It leads in the simplest way to sufficient conditions, but must be supplemented by one of the two preceding methods for an exhaustive treatment of the necessary conditions...

While Bolza (1961, pp. 104-109) used method 2 for investigating the simplest classical problem with variable endpoints, and later Bliss (1932, pp. 261-266) used the same method for the classical problem of Bolza, more recent work, e.g. [Householder (1937, pp. 485-526), Bliss (pp. 147-184), and Hestenes (1966, pp. 296-351)] have utilized the first method quoted from Bolza.

Sufficiency conditions for the problem of Bolza can be obtained by employing either method. However, the type of normality assumptions used differ from one method to the other, and the first approach apprently gained favor because it requires less stringent normality conditions. [As opposed to the second approach without modification, see for example, Bliss and Hestenes (1933, pp. 305-326) for a modification of method 2].

In this presentation, the second method was employed because of its simplicity. We were not seeking a sufficiency condition for the problem of Bolza in control notation per se. Instead a less ambitious project was investigated. We sought conditions to be

---

† Lehrbuch der Variationsrechnung Braunschweig (1900).

satisfied for a given extremal to be a local minimum with respect to endpoint variations along a prescribed endpoint manifold.

2.5 Geodetic Example

As an example of the application of the second order condition for endpoints, consider the problem of determining the minimum distance from the origin to any point on a parabola of the form

$$y = x^2 + b \tag{2.5.1}$$

In control notation the problem may be formulated as follows:

Minimize

$$J = \int_{s_o}^{s_f} ds \tag{2.5.2}$$

subject to the state variable differential constraints,

$$\frac{dx}{ds} = \cos g, \tag{2.5.3}$$

$$\frac{dy}{ds} = \sin g, \tag{2.5.4}$$

and endpoint constraints,

$$y_o = 0, \tag{2.5.5}$$

$$x_o = 0, \tag{2.5.6}$$

$$s_o = 0, \tag{2.5.7}$$

$$y_f - x_f^2 - b = 0. \tag{2.5.8}$$

The angle g is the angle between the positive x axis and a tangent to the curve. Here x and y are the state variables, g is the control variable, and s is the independent variable analogous to t in the formulation of earlier sections.

Necessary Path Conditions - The H and G functions are

$$H = \lambda_x \cos g + \lambda_y \sin g - 1 \tag{2.5.9}$$

$$G = \mu_1(y_f - x_f^2 - b) + \mu_2 x_o + \mu_3 x_o + \mu_4 s_o \tag{2.5.10}$$

The adjoint-variable Euler-Lagrange equations are

$$\dot{\lambda}_y = 0 \tag{2.5.11}$$

$$\dot{\lambda}_x = 0 \tag{2.5.12}$$

and the control-variable Euler-Lagrange equation is

$$-\lambda_x \sin g + \lambda_y \cos g = 0 \tag{2.5.13}$$

Equations (2.5.11) and (2.5.12) imply that λ_x and λ_y are constants. Solving equation (2.5.13) for the control

$$\tan g = \frac{\lambda_y}{\lambda_x} = \text{constant} \tag{2.5.14}$$

which implies

$$\sin g = \frac{\lambda_y}{\sqrt{\lambda_x^2 + \lambda_y^2}} \tag{2.5.15}$$

$$\cos g = \frac{\lambda_x}{\sqrt{\lambda_x^2 + \lambda_y^2}} \tag{2.5.16}$$

The positive sign on the radical is a consequence of the Legendre-Clebsh necessary condition (1.2.17).

Functional Relations - Integrating the state variable equations (2.5.3) and (2.5.4) with the optimal constant control g between the general initial point (x_o, y_o, s_o) and general final point (x_f, y_f, s_f) results in

$$x_f - x_o = (s_f - s_o) \cos g \tag{2.5.17}$$

$$y_f - y_o = (s_f - s_o) \sin g \tag{2.5.18}$$

Solving for the control

$$\tan g = \frac{y_f - y_o}{x_f - x_o} \tag{2.5.19}$$

Squaring both sides of equations (2.5.17) and (2.5.18) and adding yields the identity

$$(s_f - s_o)^2 = (x_f - x_o)^2 + (y_f - y_o)^2 \tag{2.5.20}$$

Solving equations (2.5.17) and (2.5.18) for the controls gives

$$\cos g = \frac{x_f - x_o}{s_f - s_o} \tag{2.5.21}$$

and

$$\sin g = \frac{y_f - y_o}{s_f - s_o} \tag{2.5.22}$$

Since the control is constant, the control is not a function of the independent variable in this case. For other problems the control may be a function of the independent variable as well as the endpoints.

Integrating the state variable equations again between the general initial point (x_o, y_o, s_o) and general intermediate point (x, y, s) and substituting the optimal control from equations (2.5.21) and (2.5.22), and rearranging yields

$$x = x_o + \frac{x_f - x_o}{s_f - s_o} (s - s_o) \tag{2.5.23}$$

$$y = y_o + \frac{y_f - y_o}{s_f - s_o} (s - s_o) \tag{2.5.24}$$

It is seen from the above equations that the state variables are clearly functions of coordinates of the initial and final state variables and of the initial and final values of the dependent variable.

The first integral of the Euler-Lagrange equations is

$$\lambda_x \cos g + \lambda_y \sin g - 1 = 0 \tag{2.5.25}$$

Solving this equation with equation (2.5.13) for λ_x and λ_y and observing equations (2.5.21) and (2.5.22) gives

$$\lambda_x = \frac{x_f - x_o}{s_f - s_o} = \frac{x_f - x_o}{\sqrt{(x_f - x_o)^2 + (y_f - y_o)^2}} \tag{2.5.26}$$

$$\lambda_y = \frac{y_f - y_o}{s_f - s_o} = \frac{y_f - y_o}{\sqrt{(x_f - x_o)^2 + (y_f - y_o)^2}} \tag{2.5.27}$$

Two forms are given above for the Lagrange multipliers as functions of endpoints; either is correct. If the second set is used, the J^* function will be independent of s_f and s_o. In either case it is clear that the Lagrange multipliers can be written as explicit functions of the coordinates of the initial and final values of the dependent variables. Equations (2.5.19), (2.5.23), (2.5.24), (2.5.26), and (2.5.27) bear out the functional dependencies hypothesized for control, state, and adjoint variables in section 2.1. Note that in deriving these equations, only path necessary conditions have been used. The transversality necessary conditions for endpoints have not been used.

<u>Necessary Endpoint Conditions</u> - The transversality conditions, (2.2.6), (2.2.7), (2.1.12), and (2.2.13) yield the following equations

$$\frac{\partial J^*}{\partial y_o} = \mu_2 - \lambda_{yo} = 0 \tag{2.5.28}$$

$$\frac{\partial J^*}{\partial x_o} = \mu_3 - \lambda_{xo} = 0 \tag{2.5.29}$$

$$\frac{\partial J^*}{\partial s_o} = \mu_4 + H_o = 0 \tag{2.5.30}$$

$$\frac{\partial J^*}{\partial y_f} = \mu_1 + \lambda_{yf} = 0 \tag{2.5.3.}$$

$$\frac{\partial J^*}{\partial x_f} = - 2\mu_1 x_f + \lambda_{xf} = 0 \tag{2.5.32}$$

$$\frac{\partial J^*}{\partial s_f} = - \lambda_{xf} \cos g_f - \lambda_{yf} \sin g_f + 1 = 0 \tag{2.5.33}$$

Since the initial point is fixed, the initial point transversality equations give no useful information.

To find the optimal endpoints, eliminate μ_1 between equations (2.5.31) and (2.5.32), yielding

$$\lambda_{xf} + 2\lambda_{yf}x_f = 0 \qquad (2.5.34)$$

Substituting λ_{xf} and λ_{yf} from equations (2.5.26) and (2.5.27) into equation (2.5.34) yields

$$\frac{x_f - x_o}{\sqrt{(x_f-x_o)^2 + (y_f-y_o)^2}} + \frac{2(y_f - y_o)\, x_f}{\sqrt{(x_f-x_o)^2 + (y_f-y_o)^2}} = 0 \qquad (2.5.35)$$

Finally, multiplying through by the radical and imposing endpoint constraints (2.5.5) and (2.5.6) gives

$$x_f(1 + 2y_f) = 0 \qquad (2.5.36)$$

The necessary conditions are satisfied if either term in the above equation is equal to zero. Solving equation (2.5.36) and equation (2.5.8) simultaneously gives the two solutions

$$x_f = \pm\sqrt{-\frac{1}{2} - b} \qquad y_f = -\frac{1}{2} \text{ (solution A)}$$

and $\qquad (2.5.37)$

$$x_f = 0 \qquad y_f = b \quad \text{(solution B)}$$

These endpoints and the corresponding multiple solutions for $b < -\frac{1}{2}$ are shown in Figure 1.3 on page 11. From the symmetry of the parabola, it is expected that either the plus or the minus sign in equation (2.5.37) will determine a solution giving the same value of distance. For this reason a distinction has not been made between the two. The necessary conditions used so far have provided no means for determining under what circumstances solution A (or solution B) is the optimum. In this case of multiple stationary solutions, the endpoint second order condition will provide a means for determining the true optimum.

Before examining the second order conditions, the paramenter μ_1 will be evaluated in terms of the general endpoints for future reference. From equations (2.5.27) and (2.5.31) it is observed that

$$\mu_1 = - \lambda_{yf} = - \frac{y_f - y_o}{\sqrt{(x_f - x_o)^2 + (y_f - y_o)^2}} \tag{2.5.38}$$

Second Order Endpoint Condition - To evaluate the endpoint sufficiency conditions, it is instructive to first determine the Φ and Ω matrices of equations (2.3.3) and (2.3.4). The constraints are

$$\psi_1: \qquad y_o = 0 \tag{2.5.39}$$

$$\psi_2: \qquad x_o = 0 \tag{2.5.40}$$

$$\psi_3: \qquad s_o = 0 \tag{2.5.41}$$

$$\psi_4: \qquad y_f - x_f^2 - b = 0 \tag{2.5.42}$$

Since there are four equations and six endpoints, there are two degrees of freedom. For conveniency let x_f and s_f be the independent variables and y_o, x_o, s_o, and y_f be the dependent variables. Then in the notation of section 2.3

$$\underline{v} = \begin{vmatrix} x_f \\ s_f \end{vmatrix}, \quad \underline{w} = \begin{vmatrix} y_o \\ x_o \\ s_o \\ y_f \end{vmatrix}, \quad \underline{r} = \begin{vmatrix} y_o \\ x_o \\ s_o \\ y_f \\ x_f \\ s_f \end{vmatrix}, \quad \underline{\psi} = \begin{vmatrix} y_o \\ x_o \\ s_o \\ y_f - x_f^2 - b \end{vmatrix} \tag{2.5.43}$$

Evaluating the matrix of partial derivatives of $\underline{\psi}$ with respect to the independent variables gives

$$\left[\frac{\partial \underline{\psi}}{\partial \underline{v}} \right] = \begin{vmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ -2x_f & 0 \end{vmatrix} \tag{2.5.44}$$

Evaluating the matrix of partial derivatives of $\underline{\psi}$ with respect to the dependent variables gives just the identity matrix

$$\frac{\partial \underline{\psi}}{\partial \underline{w}} = \begin{vmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{vmatrix} \tag{2.5.45}$$

The inverse of this matrix is obviously the identity matrix. From equations (2.5.44) and (2.4.45) the Φ matrix can be computed

$$\Phi = - \left[\frac{\partial \underline{\psi}}{\partial \underline{w}} \right]^{-1} \left[\frac{\partial \underline{\psi}}{\partial \underline{v}} \right] = \begin{vmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 2x_f & 0 \end{vmatrix} \tag{2.5.46}$$

The Ω matrix is formed by adjoining the Φ an identity matrix with the dimensions equal to the number of independent variables. In this case there are two independent variables. The Ω matrix is

$$\Omega = \left[\frac{\Phi}{I} \right] = \begin{vmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 2x_f & 0 \\ 1 & 0 \\ 0 & 1 \end{vmatrix} \tag{2.5.47}$$

With the $\left[\frac{\partial J^*}{\partial \underline{r} \partial \underline{r}} \right]$ not yet evaluated, the endpoint sufficiency condition reduces to the condition that

$$[dx_f \ ds_f] \begin{vmatrix} 4x_f^2 \dfrac{\partial^2 J^*}{\partial y_f^2} + 4x_f \dfrac{\partial^2 J^*}{\partial y_f \partial x_f} + \dfrac{\partial^2 J^*}{\partial x_f^2} & 2x_f \dfrac{\partial^2 J^*}{\partial y_f \partial s_f} + \dfrac{\partial^2 J^*}{\partial x_f \partial s_f} \\ 2x_f \dfrac{\partial^2 J^*}{\partial s_f \partial y_f} + \dfrac{\partial^2 J^*}{\partial s_f \partial x_f} & \dfrac{\partial^2 J^*}{\partial s_f^2} \end{vmatrix} \begin{vmatrix} dx_f \\ ds_f \end{vmatrix} \qquad (2.5.48)$$

must be positive definite.

If J^* can be written so that is is not a function of s_f, the sufficiency condition will be reduced to a simple inequality involving dx_f only. From the transversality equations (2.5.31) - (2.5.33) and the functional relations for λ_x and λ_y, equations (2.5.26) and (2.5.27), it is seen that this can be done.

Therefore, the sufficiency condition reduces to the condition that

$$\left[4x_f^2 \frac{\partial^2 J^*}{\partial y_f^2} + 4x_f \frac{\partial^2 J^*}{\partial y_f \partial x_f} + \frac{\partial^2 J^*}{\partial x_f^2} \right] dx_f^2 > 0 \qquad (2.5.49)$$

This result is identical to the result that would have been obtained if the fixed endpoint coordinates, y_o, x_o, and s_o had been excluded from the G function. This situation is similar to the transversality necessary conditions in that the initial points yield no information. From this example and previous experience with endpoint conditions, the following conclusion is drawn: No useful information concerning either necessary or sufficient conditions results from including in G constraints which merely fix a given endpoint coordinate.

Substituting the functional forms for λ_{xf} and λ_{yf} not involving s_f from equations (2.5.26) and (2.5.27) into the first partial derivatives of J^* with respect to y_f and x_f in equations (2.5.31) and (2.5.32) gives

$$\frac{\partial J^*}{\partial y_f} = \mu_1 + \frac{y_f - y_o}{\sqrt{(x_f-x_o)^2 + (y_f-y_o)^2}} \tag{2.5.51}$$

$$\frac{\partial J^*}{\partial x_f} = -2\mu_1 x_f + \frac{x_f - x_o}{\sqrt{(x_f-x_o)^2 + (y_f-y_o)^2}} \tag{2.5.52}$$

Forming the required partial derivatives results in

$$\frac{\partial^2 J^*}{\partial x_f^2} = -2\mu_1 + \frac{(x_f-x_o)^2 + (y_f-y_o)^2 - x_f(x_f-x_o)}{D} \tag{2.5.53}$$

$$\frac{\partial^2 J^*}{\partial y_f \partial x_f} = -\frac{(x_f-x_o)\ (y_f-y_o)}{D} \tag{2.5.54}$$

$$\frac{\partial^2 J^*}{\partial y_f^2} = \frac{(x_f-x_o)^2 + (y_f-y_o)^2 - y_f(y_f-y_o)}{D} \tag{2.5.55}$$

$$\frac{\partial^2 J^*}{\partial x_f \partial y_f} = -\frac{(y_f-y_o)\ (x_f-x_o)}{D} \tag{2.5.56}$$

where

$$D = [(x_f-x_o)^2 + (y_f-y_o)^2]^{\frac{3}{2}} \tag{2.5.57}$$

Comparing equations (2.5.54) and (2.5.56) verifies the symmetry of the $\left[\frac{\partial^2 J^*}{\partial \underline{r} \partial \underline{r}}\right]$ matrix.

Evaluating these derivatives using initial point constraint equations (2.5.5) and (2.5.6) and u_1 from equation (2.5.38) and substituting them into the endpoint sufficiency condition, equation (2.5.50) gives

$$\frac{4x_f^4}{(x_f^2 + y_f^2)^{3/2}} - 4x_f \left[\frac{x_f y_f}{(x_f^2 + y_f^2)^{3/2}} \right] + 2 \frac{y_f}{\sqrt{x_f^2 + y_f^2}} + \frac{y_f^2}{(x_f^2 + y_f^2)^{3/2}} > 0 \qquad (2.5.58)$$

for solution B ($x_f = 0$, $y_f = b$) this condition reduces to

$$\frac{1}{b} + 2 > 0 \qquad (2.5.59)$$

From the geometry in Figure 1.3, it can be seen that b is negative. The condition therefore requires that

$$0 > b > -\frac{1}{2} \qquad (2.5.60)$$

Then solution B as shown in Figure 1.3 is optimum.

For solution A ($x_f = \pm\sqrt{-\frac{1}{2} - b}$, $y_f = -\frac{1}{2}$) ,

the end sufficiency condition (2.5.58) becomes

$$\frac{4b^2 + 2b + \frac{1}{4}}{(-\frac{1}{4} - b)^{3/2}} - \frac{1}{\sqrt{-\frac{1}{4} - b}} > 0 \tag{2.5.61}$$

Combining terms yields

$$\frac{4b^2 + 3b + \frac{1}{2}}{(-\frac{1}{4} - b)^{3/2}} > 0 \tag{2.5.62}$$

In order for the denominator to be real

$$b < -\frac{1}{4} . \tag{2.5.63}$$

Under this condition inequality (2.5.62) is satisfied only if

$$b < -\frac{1}{2} . \tag{2.5.64}$$

Therefore solution A shown in Figure 1.3 is optimum for b less than $-\frac{1}{2}$. The optimal solution is summarized below.

$$x_f = 0, \quad y_f = b \qquad 0 > b > -\frac{1}{2} \tag{2.5.65}$$

$$x_f = \pm\sqrt{-\frac{1}{2} - b}\,, \quad y_f = -\frac{1}{2} \qquad b < -\frac{1}{2} \tag{2.5.66}$$

This simple example has been analyzed in great detail to emphasize the concepts developed in earlier sections and to reinforce and illustrate the notation.

2.6 A Numerical Algorithm

In order to apply the second order endpoint condition, the matrix of second partial derivatives of J^* with respect to its arguments must be determined. From equations (2.3.6) - (2.3.9) it is seen that each of these second partial derivatives is composed of two terms. The first terms, in all cases, is a second partial derivative of the function G. This derivative can be computed analytically from information given in the statement of the problem. The second term of each second partial derivative of J^* can be written in one of the following forms:

$$\frac{\partial M\big|_{t=t_f}}{\partial r_{io}} , \tag{2.6.1}$$

$$\frac{\partial M\big|_{t=t_o}}{\partial r_{if}} , \tag{2.6.2}$$

$$\frac{\partial M\big|_{t=t_f}}{\partial r_{if}} , \tag{2.6.3}$$

or

$$\frac{\partial M\big|_{t=t_o}}{\partial r_{io}} , \tag{2.6.4}$$

where M represents any of the quantities H, λ_1, λ_2, ... , λ_n and r represents any of the state variables y_i or the independent variable t.

These derivatives cannot be evaluated analytically without obtaining an analytic solution to the set of state variable and Euler-Lagrange differential equations. For most problems of practical interest in the calculus of variations, the set of nonlinear state variable and Euler-Lagrange differential equations cannot be integrated analytically. Therefore, the implementation of the endpoint sufficiency condition in most cases requires the numerical computation of partial derivatives of the forms expressed in equations (2.6.1) - (2.6.4).

Fortunately, this is not conceptually difficult for most problems in engineering which have separated end constraints. End constraints are separated if none of the endpoint constraints involves both initial values and final values; the constraints always related initial values to other initial values, or final values to other final values.

The function M evaluated at $t = t_o$ will be indicated by a subscript o:

$$M_o = M_o(y_{io}, y_{if}, t_o, t_f) \tag{2.6.5}$$

The function M evaluated at $t = t_f$ will be indicated by a subscript f:

$$M_f = M_f(y_{io}, y_{if}, t_o, t_f)$$

Before the second order condition test is applied, the problem is first solved using the necessary conditions yielding nominal endpoints y^*_{io}, y^*_{if}, t^*_o, and t^*_f and nominal Lagrange multipliers λ^*_{io} and λ^*_{if}. For brevity, let $\underline{r}^*_o$ represent a vector with elements (y^*_{1o}, y^*_{2o}, ..., y^*_{no}, t^*_o), and $\underline{r}^*_f$ represent a vector with elements (y^*_{1f}, y^*_{2f}, ..., y^*_{nf}, t^*_f), and M^* be a function evaluated with the nominal endpoints.

Numerically the derivative (2.6.1) can be approximated as

$$\frac{\partial M_f}{\partial r_{io}} = \frac{M_f(r^*_{1o}, r^*_{2o}, \ldots, r^*_{io} + \Delta, \ldots; \underline{r}^*_f) - M^*_f}{\Delta} \tag{2.6.7}$$

where Δ is a small change in the nominal initial variable r_{io}^*. If the state variable and Euler-Lagrange equations are then numerically integrated forward with the nominal Lagrange multipliers, the final nominal endpoint will not be reached. The n initial Lagrange multipliers must be adjusted in order to obtain the final nominal endpoint again. Since the n initial Lagrange multipliers give only n degrees of freedom, the nominal endpoint can be reached only if M is a function of n or less then n <u>independent</u> final values. This will be true if there is at least one equation of constraint involving the final values. With these new multipliers, the differential equations are integrated forward to the final point $\underline{r}_f^*$. M_f is then evaluated from the resulting final Lagrange multipliers and $\underline{r}_f^*$. With M_f evaluated, the desired partial derivative can be evaluated using equation (2.6.7).

The derivative (2.6.2) can be approximated numerically as

$$\frac{\partial M_o}{\partial r_{if}} = \frac{M_o(\underline{r}_o^*;\ r_{1f}^*,\ r_{2f}^*,\ \ldots;\ r_{if}^* + \Delta,\ \ldots) - M_o^*}{\Delta} \tag{2.6.8}$$

In the above equation, M_o is evaluated by making a small change in r_{if}^*, while leaving all the other values unchanged. A set of final Lagrange multipliers is then determined so that a backward numerical integration in time will yield the nominal initial values $\underline{r}_o^*$. The quantity M_o is evaluated using the resulting initial Lagrange multipliers and $\underline{r}_o^*$. With M_o computed in this manner, the desired partial derivative can be evaluated using equation (2.6.8).

The derivative (2.7.3) can be approximated numerically as

$$\frac{\partial M_f}{\partial r_{if}} = \frac{M_f(\underline{r}_o^*;\ r_{1f}^*,\ r_{2f}^*,\ \ldots,\ r_{if}^* + \Delta,\ \ldots) - M_f^*}{\Delta} \tag{2.6.9}$$

Here, M_f is evaluated by making a small change in the nominal final point coordinate r_{if}^*, while leaving all of the other final coordinates and the initial point $\underline{r}_o^*$ unchanged. A set of initial Lagrange multipliers is then determined so that a forward integration from the nominal initial point will yield the varied final point $(r_{1f}^*, r_{2f}^*, \ldots, r_{if}^* + \Delta, \ldots)$. The forward integration is then performed to the varied final point, and M_f is evaluated using the resulting final Lagrange multipliers and the coordinates of the varied final point.

The final derivative (2.6.4) can be approximated numerically as

$$\frac{\partial M_o}{\partial r_{io}} = \frac{M_o(r_{1o}^*, r_{2o}^*, \ldots, r_{io}^* + \Delta, \ldots; \underline{r}_f^*) - M_o^*}{\Delta} \tag{2.6.10}$$

Here, M_o is evaluated by making a small change in the nominal initial point r_{io}^*, while leaving all of the other initial coordinates and the final point $\underline{r}_f^*$ unchanged. A set of final Lagrange multipliers is then determined so that a backward integration in time from the nominal final point will yield the varied initial point $(r_{1o}^*, r_{2o}^*, \ldots, r_{io}^* + \Delta, \ldots)$. The backward integration is then performed to the varied initial point, and M_o is evaluated using the resulting initial Lagrange multipliers and the coordinates of the varied initial point.

Using the above techniques, the matrix of second partial derivatives of J^* with respect to its arguments can be evaluated. Because of the identities (2.3.11) - (2.3.14), there is some choice as to which of the above derivatives is used to evaluate the sufficiency condition. It is a simple matter to numerically evaluate the matrix Ω from the nominal initial and final points and to test the matrix $\Omega^T \dfrac{\partial^2 J^*}{\partial \underline{r} \partial \underline{r}} \Omega$ for positive-definiteness.

CONCLUSIONS

It has been shown that once the first necessary path conditions have been applied, a calculus of variations problem with variable endpoints is reduced to a problem of the minimization of a function of several variables.

Analytical application of the second order endpoint condition requires the analytical integration of the set of state variable and Euler-Lagrange differential equations. Since in most cases this is difficult or impossible, the algorithm developed for the numerical implementation of the second order endpoint test should be an effective computational tool in complex applications. For example, through the use of the second order endpoint test, a complete class of nonoptimal solutions can be discarded immediately upon encounter. Without the aid of the second order endpoint test an investigator would have no indication that solutions he is generating are non-optimal whether he encounters multiple solutions or not.

It could be argued, when multiple stationary solutions are obtained, that a comparison of solutions would quickly yield which one was optimal. However such a comparison technique fails if it is not known apriori exactly how many multiple solutions exist. One has no criteria in general for determining in advance just how many multiple stationary solutions a problem may have, so that a direct comparison technique is unreliable unless every multiple solution is somehow found.

APPENDIX A

MINIMIZATION OF A FUNCTION OF SEVERAL VARIABLES

In the proof that follows, free use will be made of the notation and conventions established at the beginning of section 2.3. Following the methods of Vincent and Cliff (1970), consider the problem of minimizing a function of several variables

$$J = J(\underline{w}, \underline{v}) \tag{A.1}$$

subject to the constraints

$$\underline{\psi}(\underline{w}, \underline{v}) = \underline{0} \tag{A.2}$$

where both J and $\underline{\psi}$ are functions of class C^2 and the constraints are such that the determinant of the Jacobian

$$\left[\frac{\partial \underline{\psi}}{\partial \underline{v}} \right] \tag{A.3}$$

is nonsingular. The dimension of $\underline{\psi}$ and $\underline{w}$ is assumed to be p and the dimension of $\underline{v}$ is assumed to be q.

A.1 Method of Implicit Functions

Since $\underline{\psi}$ is of C^2 and condition (A.3) has been postulated, the implicit function theorem (Buck, 1965, pp. 283-286) states that equation (A.2) implicitly assures the existence of the vector function $\underline{W}$ explicitly relating the dependent variables $\underline{w}$ to the independent variables $\underline{v}$

$$\underline{w} = \underline{W}(\underline{v}) = \begin{vmatrix} W_1(\underline{v}) \\ W_2(\underline{v}) \\ \vdots \\ W_p(\underline{v}) \end{vmatrix} \tag{A.4}$$

By substituting (A.4) into (A.1), J becomes a function of $\underline{v}$ only

$$J = J(W(\underline{v}), \underline{v}) \tag{A.5}$$

Define the general value of independent variables $\underline{v}$ in a small neighborhood of an optional point $\underline{v}^o$

$$\underline{v} = v^o + \epsilon \ \underline{c} \tag{A.6}$$

where $\underline{c}$ is a vector of arbitrarily chosen small, but non-zero, constants and ϵ is a scalar multiplier. Then, from (A.1) and (A.2)

$$J = J\ [\underline{W}(\underline{v}^o + \epsilon\ \underline{c}),\ \underline{v}^o + \epsilon\ \underline{c}] \tag{A.7}$$

$$\psi\ [\underline{W}(\underline{v}^o + \epsilon\ \underline{c}),\ v^o + \epsilon\ \underline{c}] = 0 \tag{A.8}$$

Now J is a function of ϵ only, and the necessary condition for an ordinary local extremum is

$$\frac{dJ}{d\epsilon} = \frac{\partial J^T}{\partial \underline{w}}\ \underline{h} + \frac{\partial J^T}{\partial \underline{v}}\ \underline{c} = 0 \tag{A.9}$$

where $\underline{h}$ is the vector with elements $h_i = \dfrac{\partial W_i}{\partial \epsilon}$. The vector $\underline{h}$ represents changes in the dependent variables $\underline{w}$ corresponding to the changes $\underline{c}$ in the independent variables. Differentiating equation (A.8) with respect to ϵ yields

$$\left[\frac{\partial \psi}{\partial \underline{w}}\right] \underline{h} + \left[\frac{\partial \psi}{\partial \underline{v}}\right] \underline{c} = 0 \tag{A.10}$$

Solving for $\underline{h}$ yields

$$\underline{h} = - \left[\frac{\partial \psi}{\partial \underline{w}}\right]^{-1} \left[\frac{\partial \psi}{\partial \underline{v}}\right] \underline{c} \tag{A.11}$$

Substituting (A.11) into (A.9) and rearranging gives

$$\frac{dJ}{d\epsilon} = \left[\frac{\partial J}{\partial \underline{v}}^T - \frac{\partial J}{\partial \underline{w}}^T \left[\frac{\partial \underline{\underline{\psi}}}{\partial \underline{w}}\right]^{-1} \left[\frac{\partial \underline{\underline{\psi}}}{\partial \underline{v}}\right] \right] \underline{c} = \underline{0} \tag{A.12}$$

Since $\underline{c}$ is an arbitrary nonzero vector, each of the elements of the vector in parenthesis must be equal to $\underline{0}$ at the optimal point.

$$\frac{\partial J}{\partial \underline{v}}^T - \frac{\partial J}{\partial \underline{w}}^T \left[\frac{\partial \underline{\underline{\psi}}}{\partial \underline{w}}\right]^{-1} \left[\frac{\partial \underline{\underline{\psi}}}{\partial \underline{v}}\right] = \underline{0} \tag{A.13}$$

If equation (A.13) is satisfied, then a further necessary condition for a local interior minimum is that

$$\left.\frac{d^2 J}{d\epsilon^2}\right|_{v_o} \tag{A.14}$$ †

must be positive semi-definite for arbitrary values of $\underline{h}$ and $\underline{c}$ satisfying equation (A.11).

Before evaluating this expression, an identity for taking the partial derivative of the inverse of a matrix must be developed. Let A_{ij} represent a general element of the $\left[\frac{\partial \underline{\underline{\psi}}}{\partial \underline{w}}\right]^{-1}$ matrix:

$$A_{ij} = \left[\frac{\partial \underline{\underline{\psi}}}{\partial \underline{w}}\right]^{-1}_{ij} \tag{A.15}$$

Then in indicial notation the definition of inverse may be expressed as

$$\delta_{qj} = \frac{\partial \psi_q}{\partial w_m} A_{mj} \tag{A.16}$$

---

† Equation (A.13) and (A.14) positive definite is sufficient for a local interior minimum.

where δ_{qj} is the Kroniker delta. Premultiplying by A_{iq} gives

$$A_{ij} = A_{iq} \frac{\partial\psi_q}{\partial w_m} A_{mj} \tag{A.17}$$

Taking the partial derivative of both sides yields

$$\frac{\partial}{\partial r_n}(A_{ij}) = \frac{\partial}{\partial r_n}(A_{iq}) \delta_{qj} + A_{iq} \frac{\partial}{\partial r_n}\left(\frac{\partial\psi_q}{\partial w_m}\right) A_{mj}$$
$$+ \delta_{im} \frac{\partial}{\partial r_n}(A_{mj}) \tag{A.18}$$

Since δ_{ij} represents a constant, equation (A.18) reduces to

$$\frac{\partial}{\partial r_n}(A_{ij}) = \frac{\partial}{\partial r_n}(A_{ij}) + A_{iq} \frac{\partial}{\partial r_n}\left(\frac{\partial\psi_q}{\partial w_m}\right) A_{mj} + \frac{\partial}{\partial r_n}(A_{ij}) \tag{A.19}$$

which gives the desired identity:

$$\frac{\partial}{\partial r_n}(A_{ij}) = -A_{iq} \frac{\partial}{\partial r_n}\left(\frac{\partial\psi_q}{\partial w_m}\right) A_{mj} \tag{A.20}$$

From this point on results must be expressed in indicial notation since $\frac{\partial}{\partial r_n}\left(\frac{\partial\psi_q}{\partial w_m}\right)$ is a tensor. In indicial notation equation (A.12) becomes

$$\frac{dJ}{d\varepsilon} = \left[\frac{\partial J}{\partial v_k} - \frac{\partial J}{\partial w_i} A_{ij} \frac{\partial\psi_j}{\partial v_k}\right] c_k \tag{A.21}$$

Using equations (A.20) and (A.21) the second order condition (A.14) becomes

$$\frac{d^2J}{d\epsilon^2} = \frac{\partial^2 J}{\partial w_i \partial v_k} c_k h_i + \frac{\partial^2 J}{\partial v_n \partial v_k} c_k c_n$$

$$- \frac{\partial^2 J}{\partial w_n \partial w_i} A_{ij} \frac{\partial \psi_j}{\partial v_k} c_k h_n - \frac{\partial^2 J}{\partial v_n \partial w_i} A_{ij} \frac{\partial \psi_j}{\partial v_k} c_k c_n$$

$$+ \frac{\partial J}{\partial w_i} A_{iq} \frac{\partial^2 \psi_q}{\partial w_n \partial w_m} A_{mj} \frac{\partial \psi_j}{\partial v_k} c_k h_n \tag{A.22}$$

$$+ \frac{\partial J}{\partial w_i} A_{iq} \frac{\partial^2 \psi_q}{\partial v_n \partial w_m} A_{mj} \frac{\partial \psi_j}{\partial v_k} c_k c_n$$

$$- \frac{\partial J}{\partial w_i} A_{ij} \frac{\partial^2 \psi_j}{\partial w_p \partial v_k} c_k h_p$$

$$- \frac{\partial J}{\partial w_i} A_{ij} \frac{\partial^2 \psi_j}{\partial v_n \partial v_k} c_k c_n$$

The indicial notation representation of equation (A.11)

$$h_m = - A_{mj} \frac{\partial \psi_j}{\partial v_k} c_k \tag{A.23}$$

appears in four terms of equation (A.22). Regrouping terms, (A.23) can be written

$$\frac{d^2J}{d\epsilon^2} = \left(\frac{\partial^2 J}{\partial v_n \partial v_k} - \frac{\partial J}{\partial w_i} A_{ij} \frac{\partial^2 \psi_j}{\partial v_n \partial v_k}\right) c_k c_n$$

$$+ \left(\frac{\partial^2 J}{\partial w_i \partial v_k} - \frac{\partial J}{\partial w_p} A_{pj} \frac{\partial^2 \psi_j}{\partial w_i \partial v_k}\right) c_k h_i \tag{A.24}$$

$$+ \left(\frac{\partial^2 J}{\partial v_n \partial w_i} - \frac{\partial J}{\partial w_p} A_{pq} \frac{\partial^2 \psi_q}{\partial v_n \partial w_i}\right) h_i c_n \tag{A.24}$$

$$+ \left(\frac{\partial^2 J}{\partial w_j \partial w_i} - \frac{\partial J}{\partial w_p} A_{pq} \frac{\partial^2 \psi_q}{\partial w_i \partial w_j}\right) h_i h_j$$

Equation (A.24) and equation (A.13) provide a set of first and second order conditions for $J(\underline{w}, \underline{v})$ to be minimum subject to the constraints $\underline{\psi}(\underline{w}, \underline{v}) = 0$.

A.2 Method of Lagrange Multipliers

The first and second order conditions can be put in a form which is more convenient to use by defining the augmented function

$$J^*(\underline{w}, \underline{v}, \underline{\mu}) = J(\underline{w}, \underline{v}) + \underline{\mu}\, \underline{\psi}(\underline{w}, \underline{v}) \tag{A.25}$$

where $\underline{\mu}$ is a vector of constant multipliers called Lagrange multipliers.

If the Lagrange multipliers are given by the identity

$$\mu_j = -\frac{\partial J}{\partial w_i} A_{ij} \tag{A.26}$$

or in matrix notation

$$\underline{\mu} = -\frac{\partial J}{\partial \underline{w}}^T \left[\frac{\partial \underline{\psi}}{\partial \underline{w}}\right]^{-1} \tag{A.27}$$

several observations can be made. Thus necessary conditions for J to be a minimum (A.13) become

$$\frac{\partial J^*}{\partial \underline{v}} = \underline{0} \tag{A.28}$$

The second order conditions (A.24) may be written as

$$\left(\frac{\partial^2 J}{\partial v_n \partial v_k} + \mu_j \frac{\partial^2 \psi_j}{\partial v_n \partial v_k}\right) c_k c_n + \left(\frac{\partial^2 J}{\partial w_i \partial v_k} + \mu_j \frac{\partial^2 \psi_j}{\partial w_i \partial v_k}\right) c_k h_i$$
$$+ \left(\frac{\partial^2 J}{\partial v_n \partial w_k} + \mu_j \frac{\partial^2 \psi_j}{\partial v_n \partial w_i}\right) h_i c_n + \left(\frac{\partial^2 J}{\partial h_j \partial h_i} + \mu_q \frac{\partial^2 \psi_q}{\partial w_i \partial w_j}\right) h_i h_j \tag{A.29}$$

Define the vectors $\underline{r}$ and $\underline{d}$ as

$$\underline{r} = \begin{vmatrix} w_1 \\ w_2 \\ \vdots \\ w_p \\ v_1 \\ v_2 \\ \vdots \\ v_q \end{vmatrix} \qquad \underline{d} = \begin{vmatrix} h_1 \\ h_2 \\ \vdots \\ h_p \\ c_1 \\ c_2 \\ \vdots \\ c_q \end{vmatrix} \tag{A.30}$$

The second order condition (A.29) may now be put in compact matrix notation by using equation (A.25) and definitions (A.30):

$$\underline{d}^T \left[\frac{\partial^2 J^*}{\partial \underline{r} \partial \underline{r}} \right] \underline{d} \tag{A.31}$$

To guarantee that $\underline{c}$ and $\underline{h}$ in vector $\underline{d}$ satisfy equation (A.11), elements of $\underline{d}$ may be expressed as functions of the independent constants $\underline{c}$ only. In matrix notation this may be expressed as

$$\underline{d} = \begin{bmatrix} \bar{\phi} \\ \hline I \end{bmatrix} \underline{c} = \Omega \, \underline{c} \tag{A.32}$$

where

$$\underline{\Phi} = -\left[\frac{\partial \underline{\psi}}{\partial \underline{w}}\right]^{-1}\left[\frac{\partial \underline{\psi}}{\partial \underline{v}}\right] \tag{A.33}$$

and I is a q by q indentity matrix. The second order condition (A.32) may then be written as

$$\underline{c}^T \Omega^T \left[\frac{\partial^2 J^*}{\partial \underline{r}\, \partial \underline{r}}\right] \Omega\, \underline{c} \tag{A.34}$$

The advantage of the Lagrange multiplier technique is that the first and second order conditions can be expressed in a compact matrix notation.

REFERENCES

Bliss, G. A., Lectures on the Calculus of Variations, University of Chicago Press, Chicago (1946).

Bliss, G. A. and Hestenes, M. R., Sufficient Conditions for a Problem of Mayer in the Calculus of Variations, Transactions of the Am. Math. Soc., Vol. 35, 1933.

Bliss, G. A., The Problem of Bolza in the Calculus of Variations, Annals of Mathematics, Vol. 33, 1932.

Bolza, O., Lectures on the Calculus of Variations, Dover Publications, New York (1961); also, University of Chicago Press, Chicago (1904).

Bryson, A. E., and Ho, Y. C., Optimization Estimation, and Control, Blaisdell, New York (1969).

Buck, R. C., Advanced Calculus, McGraw-Hill Book Co., New York (1965).

Hestenes, M. R., Calculus of Variations and Optimal Control Theory, Wiley and Sons, New York (1966).

Hildebrand, F. B., Advanced Calculus for Applications, Prentice-Hall, Englewood Cliffs, N. J. (1948).

Householder, A. S., "The Dependency of a Focal Point Upon Curvature in the Calculus of Variations," Contributions to the Calculus of Variations, University of Chicago Press, Chicago (1937), pp. 485-526.

Johnson, D. C., and Gibson, J. E., "Singular Solutions in Problems of Optimal Control," IEEE Transactions on Automatic Control, Vol. AC-8 (1963), pp. 4-14.

Kelly, H. J., Kopp, R. E., and Moyer, H. G., "Singular Extremals," Topics in Optimization, edited by Leitmann, G., Academic Press, New York (1967).

Leitmann, G., An Introduction to Optimal Control, McGraw-Hill Book Co., New York (1966), pp. 57-8.

Pontryagin, L.S., Boltyanskii, V. G., Gamkrelidze, R.V., and Mischenko, E.F., The Mathematical Theory of Optimal Processes, edited by Neustadt, L. W., translated by Trirogoff, K. M., Wiley and Sons, New York (1962).

Robbins, H. M., "Optimality of Intermediate Thrust Arcs of Rocket Trajectories," American Institute of Aeronautics and Astronautics Journal, Vol. 3, No. 6 (1965), pp. 1094-8.

Vincent, T. L., and Brusch, R. G., "Minimum Time Aircraft Trajectories between Two Points in Range Altitude Space," NASA CR-631, National Aeronautics and Space Administration, Washington, D. C. (1966).

Vincent, T. L. and Cliff, E. M., "Maximum-Minimum Sufficiency and Lagrange Multipliers," AIAA Journal, January 1970.