H. M. Robbins

A Generalized Legendre-Clebsch Condition for the Singular Cases of Optimal Control

Abstract: For certain optimal control problems, some of the extremal trajectories generated by simultaneous solution of the state and adjoint equations may include arcs of a special character, called "singular" arcs. The optimality of singular arcs has been the subject of considerable uncertainty, since the classical criteria are inapplicable or inconclusive. This uncertainty has recently been reduced by the discovery of additional necessary conditions for the optimality of singular arcs. The principal result of this paper is a general statement and proof of these conditions, in the form of a "generalized Legendre-Clebsch condition" which reduces to the classical Legendre-Clebsch condition when applied to nonsingular arcs, and gives additional necessary conditions when applied to singular arcs. Other results include a classification of the possible singular arcs, a useful extension of the conventional optimal-control formalism (by the introduction of "generalized Hamiltonians" and "generalized control transformations"), and some interesting variational formulae.

1. Introduction

For certain optimal control problems, some of the extremal trajectories generated by simultaneous solution of the equations of motion and the Euler-Lagrange equations may include arcs of a special character, called "singular" arcs. The distinction between singular and nonsingular arcs of extremals goes back to classical studies in the calculus of variations, but singular arcs have not been extensively studied until quite recently. A general definition (slightly modernized) may be stated as follows: An extremal arc is singular if, at each point of the arc, there is some allowable first-order weak control variation which leaves the value of the variational Hamiltonian unchanged to second order.

When the control vector is in the interior of its allowed region, and the necessary partial derivatives exist, this definition can be restated as follows: An extremal arc is singular if the matrix $\mathbf{H}_{uu}$ (whose elements are second partial derivatives of the variational Hamiltonian with respect to components of the control vector $\mathbf{u}$) is singular everywhere on the arc, i.e., if its determinant vanishes identically on the arc.

The most common and most important cases involving singular arcs arise when the variational Hamiltonian depends *linearly* on some control variable, with a coefficient that vanishes identically on the arc. In these cases there are finite control variations which do not affect the value of the Hamiltonian, so the Pontriagin maximum principle (or its classical counterpart, the Weierstrass condition) does not directly determine a unique optimal control as a function of the state and adjoint variables. Instead, the optimal control is determined indirectly by the requirement that the coefficients of the linearly-appearing control variables remain zero for the duration of the singular arc. This is discussed at length in Section 4.

Singular arcs of the "linear" type have been extensively studied in recent years, but they are not the only type of singular arc possible. There also exist singular arcs where all the control variables occur nonlinearly in the Hamiltonian and are uniquely determined by the maximum principle. Singular arcs of the nonlinear type (or of mixed type) have not been extensively studied. However, examples are easy to construct, and at least one example has arisen in an important practical problem. A geometric interpretation of the various types of singular arcs (linear, nonlinear, and mixed) will be given in Section 3. Singular arcs of the nonlinear type will be discussed again in Section 8.

The optimality of extremals which include singular arcs has been the subject of considerable doubt and discussion, since the classical criteria are either inapplicable to singular arcs, or inconclusive. The difficulty is associated with the classical Legendre-Clebsch condition. This condition has two forms: a weak form and a strong form. Stated in modern terms (i.e., in a form compatible with Pontriagin's maximum principle) the weak Legendre-Clebsch condition requires the matrix $\mathbf{H}_{uu}$ to be negative semidefinite. It is an immediate corollary of the maximum principle (when the control vector is in the interior of its allowed region) and is a classical necessary condition for optimality. The strong form of the Legendre-Clebsch condition requires $\mathbf{H}_{uu}$ to be negative definite. For nonsingular arcs, the two forms are equivalent, but on singular arcs only the weak form holds. The classical theorems giving *sufficient* conditions for optimality assume the strong form of the Legendre-Clebsch condition, and are therefore inapplicable to extremals with singular arcs. In particular, the classical Jacobi theory, which leads to the Jacobi necessary condition (that the extremal must not include a pair of conjugate points) and to a sufficient condition for local optimality, does not apply to singular arcs.

It is easy to see that for singular arcs, the weak Legendre-Clebsch condition must be supplemented by additional necessary conditions. For example, consider the problem of minimizing the integral of $ay^2 + b\dot{y}^2$ from 0 to T, with $y = 0$ at the endpoints. This problem has no solution unless $b \geq 0$ (the weak Legendre-Clebsch condition). If $b > 0$, the extremal $y \equiv 0$ is nonsingular and the Jacobi theory may be used to prove optimality for values of T such that $aT^2 + b\pi^2 > 0$. If $b = 0$, this extremal is singular, and its optimality depends on the additional necessary condition $a \geq 0$.

The principal purpose of the present paper is to state and prove a generalized Legendre-Clebsch condition which for nonsingular arcs reduces to the classical Legendre-Clebsch condition, and for singular arcs gives additional necessary conditions like the one in the example above. It seems probable that when the generalized Legendre-Clebsch condition is satisfied, a generalization of the Jacobi theory should be possible, giving a Jacobi condition for extremals with singular arcs, and an assurance of local optimality if this condition is strongly satisfied, i.e., if the extremal (including its end-points) does not contain a pair of conjugate points. However, this is beyond the scope of the present paper.

The generalized Legendre-Clebsch condition may be regarded as an extension and generalization of the work of Kelley.[1] Considering a problem with only a single control variable, so $\mathbf{H}_u$ and $\mathbf{H}_{uu}$ reduce to scalars, Kelley showed that if $\mathbf{H}_{uu} = 0$, an additional necessary condition is (in a form suggested by Bryson)

$$\frac{\partial}{\partial \mathbf{u}} [(d/dt)^2 \mathbf{H}_u] \geq 0. \tag{1}$$

(Note: Kelley's own statement of this condition has the inequality sign reversed. This is because Kelley uses a min-H formulation of the optimality conditions instead of the max-H formulation adopted in the present paper.) When Kelley's test quantity is zero, additional conditions are needed. Kopp and Moyer (private communication) conjectured that a general necessary condition for the case of a single control variable is that the first nonvanishing member of the sequence $\{\mathbf{q}_m\}$, where

$$\mathbf{q}_m = \frac{\partial}{\partial \mathbf{u}} [(d/dt)^m \mathbf{H}_u], \qquad m = 0, 1, 2, \cdots, \tag{2}$$

must occur for *even* $m (m = 2k)$ and have the correct sign. Subsequent work by myself, and independently by Kopp and Moyer,[2,3] has shown this conjecture to be correct, except that the evenness condition is always automatically satisfied and hence need not be considered. The correct sign turns out to be $(-1)^{k+1}$, so the generalized necessary condition for cases with a single control variable is

$$(-1)^k \frac{\partial}{\partial \mathbf{u}} [(d/dt)^{2k} \mathbf{H}_u] \leq 0. \tag{3}$$

This evidently reduces to the Legendre-Clebsch condition if $k = 0$, and to the Kelley condition if $k = 1$. It is convenient to call $2k$ the *degree* of singularity of the arc. Then nonsingular arcs are singular of degree zero, and arcs for which Kelley's condition holds strongly are singular of degree two. The intermediate-thrust arcs of rocket trajectories in vacuum are singular of degree four (Robbins[4]).

The conditions derived for one control variable can be readily generalized to cases with multi-component control vectors. This gives the generalized Legendre-Clebsch condition in its most complete form. In multi-component cases, arcs with an *odd* degree of singularity can occur, but are necessarily non-optimal. It is interesting to note that this result is consonant with the original Kopp-Moyer conjecture.

The generalized Legendre-Clebsch condition is not the only approach to the optimality problem for singular arcs. Alternative approaches are the Green's-theorem method of Miele[5] and the transformation methods of Kelley[6] and Goh.[7] The transformation methods are about as powerful as the method described here, but are more laborious to apply. Miele's method is the best when it is applicable, since it can sometimes establish absolute optimality. However, it is usually inapplicable. A comprehensive discussion of singular arcs, with many original results and an excellent list of references, has recently been given by Johnson.[8]

2. Formulation of the problem

The optimization problem is assumed to be stated in the Mayer-Pontriagin form. That is, a system of first-order, ordinary differential equations (state equations) is given in the form

$$\dot{x}_i = f_i(\mathbf{x}, \mathbf{u}, t), \qquad i = 1, 2, \cdots n, \tag{4}$$

where $\dot{x}_i$ is the i^{th} component of the n-component state vector $\mathbf{x}$, and $\mathbf{u}$ is a control vector of dimensionality $n_c \geq n$. The control vector is subject to some given set of state-independent control constraints. These may be compactly expressed in the form $\mathbf{u} \in U$, where U is some given region in the control space. The assumption that the control constraints are state-independent is not very restrictive, since state-dependent constraints can generally be restated in a state-independent form. It could readily be eliminated, but this would complicate the discussion. The optimization problem is to find a trajectory (i.e., a solution of (4) consistent with the control constraints) that satisfiies given boundary conditions at an initial time t_0 and a final time t_f, and minimizes a given function of $\mathbf{x}(t_f)$. The first-order necessary conditions for optimality are given by the Pontriagin maximum principle

$$H(\mathbf{x}, \mathbf{p}, \mathbf{u}, t) = \max_{\mathbf{u}^* \epsilon U} H(\mathbf{x}, \mathbf{p}, \mathbf{u}^*, t), \tag{5}$$

where H is the Hamiltonian function

$$H(\mathbf{x}, \mathbf{p}, \mathbf{u}, t) = \sum_{i=1}^{n} p_i f_i(\mathbf{x}, \mathbf{u}, t) = \mathbf{p}^T \mathbf{f} \tag{6}$$

and $\mathbf{p}$, which is an n-component adjoint vector that is a continuous function of time, obeys the differential equation

$$\dot{p}_j = -\partial H / \partial x_j, \qquad j = 1, 2, \cdots n, \tag{7}$$

and satisfies certain transversality conditions at times t_0 and t_f. It will be convenient to introduce a compact vector-matrix notation in which partial differentiation with respect to components of a vector is denoted by using that vector as a subscript In this notation, (4) and (7) can be written in the familiar canonical form

$$\dot{\mathbf{x}} = \mathbf{H}_p \tag{8a}$$

$$\dot{\mathbf{p}} = -\mathbf{H}_x. \tag{8b}$$

We shall restrict attention to arcs in which $\mathbf{u}$ is in the interior of its allowed region U, either originally or as the result of a control-reduction process. This means that if the optimal $\mathbf{u}$ is on the boundary of U, it is generally possible to use this fact to eliminate one or more control variables, thereby defining a new $\mathbf{u}$ of reduced dimensionality, and a corresponding new U in a new control space (also of reduced dimensionality) such that the new $\mathbf{u}$ is in the interior of the new U. For arcs of the chosen type, and assuming existence and continuity of the relevant partial derivatives, (5) implies the Euler-Lagrange condition

$$\mathbf{H}_u = 0 \tag{9}$$

and the weak Legendre-Clebsch condition that the matrix $\mathbf{H}_{uu}$ be negative semidefinite:

$$\mathbf{H}_{uu} \leq 0. \tag{10}$$

If the matrix $\mathbf{H}_{uu}$ has r null eigenvectors, the arc will be said to be singular of rank r, and to have r "singular controls" and $n_c - r$ "nonsingular controls." (*Rank* of singularity, as defined here for multivariable cases, should not be confused with *degree* of singularity, defined in the previous section in connection with one singular control variable.)

The nature and significance of singular arcs can be made clearer by using a well-known geometrical interpretation of Eqs. (4)–(7). If $\mathbf{u}$ is varied over its allowed region U while $\mathbf{x}$ and t are held fixed, the vector $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \mathbf{u}, t)$ traces out a pointset $\mathbf{S}(\mathbf{x}, t)$ in a rate-space called the "hodograph space." In terms of this pointset, (4) and the control constraints can be reduced to the single relation

$$\dot{\mathbf{x}} \,\epsilon\, \mathbf{S}(\mathbf{x}, t). \tag{11}$$

If the pointset S is not convex, the optimization problem may not have a solution with piecewise-continuous controls, since a "chattering" control (rapid switching among two or more values of $\mathbf{u}$) may be able to generate an average $\dot{\mathbf{x}}$ which is desirable for optimality but not realizable by any single allowed value of $\mathbf{u}$. In such cases, the standard procedure is to consider a "relaxed" optimization problem in which S is augmented to make it convex. The solution of the relaxed problem can be directly translated into a solution of the orginal problem. To simplify the discussions below, S will be assumed convex. This does not entail any significant loss of generality.

If we introduce the abbreviation $\mathbf{f}^* = \mathbf{f}(\mathbf{x}, \mathbf{u}^*, t)$ where $\mathbf{u}^*$ is a possibly non-optimal control, (5) and (6) state that the optimal $\dot{\mathbf{x}}$, corresponding to the optimal control, must satisfy the relation

$$\mathbf{p}^T \dot{\mathbf{x}} = \max_{\mathbf{f}^* \epsilon \mathbf{S}} (\mathbf{p}^T \mathbf{f}^*). \tag{12}$$

The geometrical interpretation of this relation is that there is a hyperplane **P** in the hodograph space which has the adjoint vector $\mathbf{p}$ as its positive normal, and "rests" on the pointset S. That is, S and **P** have at least one point in common, and all points of S lie either on **P** or in the negative halfspace defined by **P**. The optimal $\dot{\mathbf{x}}$ (and hence, the optimal control) corresponds to a point of contact between S and **P**, which is a point of tangency if S possesses a smoothly turning normal in the neighborhood of the contact. In general, both S and **P** change with time, so the hyperplane rolls on a moving surface. If S and **P** have only one point of contact, the optimal control and the instantaneous rates of change of S and **P** are uniquely determined. The contact may be non-unique at isolated instants of time (this gives one type of *corner*) or throughout one or more finite-duration arcs of the trajectory.

Let $\mathbf{S}_P$ denote the set of contact points, i.e., the intersection of S and **P**. The convexity assumption for S implies that $\mathbf{S}_P$ also is convex, so $\mathbf{S}_P$ consists of a single point, or a segment of a straight line, or a convex region in a two-dimensional plane, or more generally a convex region in a linear subspace of the hodograph space, of dimensionality r'. From the geometrical viewpoint described above, the control variables are merely a convenient parameterization of the pointset $\mathbf{S}(\mathbf{x}, t)$. By a transformation of the control variables if necessary, it is always possible to ensure that $\mathbf{S}_P$ is *linearly* parameterized. Then for $\mathbf{f} \,\epsilon\, \mathbf{S}_P$ there are r' components of $\mathbf{u}$ that appear linearly in the state equations and do not affect the value of the Hamiltonian, so their values are not directly determined by the maximum principle. These r' components are singular control variables of the "linear" type mentioned in the Introduction. The remaining $r - r'$ singular control variables occur nonlinearly in the Hamiltonian, and are directly and uniquely determined by the maximum principle.

If the matrix $\mathbf{H}_{pu}$ has full rank (i.e., if each control variable really represents an independent degree of control freedom for weak perturbations) the strong Legendre-Clebsch condition $\mathbf{H}_{uu} < 0$ has an obvious geometrical interpretation: it states that contact between $\mathbf{S}$ and $\mathbf{P}$ is unique, and is at most of second order. Non-unique contact implies the existence of at least one "linear" singular control variable; unique but higher-order contact means singular control variables of the "nonlinear" type.

3. Generalized variational hamiltonians

Before proceeding further with the discussion of singular arcs, it is expedient to introduce the concept of a "generalized Hamiltonian." This concept greatly simplifies the treatment of complex singular cases, but its utility is by no means restricted to such cases. It appears to be a convenient general tool for the optimal-control theorist as well as for the user of optimal control theory.

Consider the optimization problem stated in Section 2, and the corresponding Hamiltonian given by (6). Now suppose that as a direct or indirect consequence of the maximum principle, it can be shown that the optimal control vector $\mathbf{u}$ can be expressed as a function of $\mathbf{x}$, $\mathbf{p}$, t, and a new control vector $\mathbf{v}$ (of smaller dimensionality) which remains to be determined. That is,

$$\mathbf{u} = \mathbf{u}(\mathbf{x}, \mathbf{p}, t, \mathbf{v}). \tag{13}$$

Since rescaling the adjoint vector $\mathbf{p}$ by an arbitrary positive constant factor cannot change the optimal control, the function $\mathbf{u}(\mathbf{x}, \mathbf{p}, t, \mathbf{v})$ is necessarily a homogeneous function of the components of $\mathbf{p}$, of degree zero. We shall restrict attention to cases in which $\mathbf{u}$ and $\mathbf{v}$ are in the interiors of their allowed regions, so constraints need not be considered explicitly. Substitution of (13) into the Hamiltonian gives a new Hamiltonian:

$$H_{\text{new}}(\mathbf{x}, \mathbf{p}, \mathbf{v}, t) = H_{\text{old}}[\mathbf{x}, \mathbf{p}, \mathbf{u}(\mathbf{x}, \mathbf{p}, t, \mathbf{v}), t]. \tag{14}$$

The new (generalized) Hamiltonian will not in general be of the form shown in (6), that is, it will not in general be linear in the adjoint vector $\mathbf{p}$. However, it will be a homogeneous function of the components of $\mathbf{p}$, of degree one. It therefore satisfies the Euler identity for such functions, namely

$$H \equiv \mathbf{p}^T \mathbf{H}_p. \tag{15}$$

Exactly the same equations and trajectories may be obtained from the generalized Hamiltonian [using (8) and the maximum principle] as would be obtained by using the original Hamiltonian and substituting (13) into the results. This equivalence is immediately evident for trajectory arcs in which $\mathbf{u}$ and $\mathbf{v}$ are in the interiors of their allowed regions: the equation $\mathbf{H}_u = 0$ for the old Hamiltonian implies $\mathbf{H}_v = 0$ for the new, and also implies that the values of $\dot{\mathbf{x}} = \mathbf{H}_p$ and $\dot{\mathbf{p}} = -\mathbf{H}_x$ do not depend on which Hamiltonian is used. The equivalence can also be shown for arcs in which $\mathbf{u}$ and $\mathbf{v}$ are influenced by constraints that are constant or depend on t only, though this is unnecessary for present purposes.

All, or nearly all, of the mathematical formalism that has been developed for conventional variational Hamiltonians can be extended to the generalized variational Hamiltonians described in this section. Where derivations make use of (6), they can be modified to use (15) instead. For example, some elegant second-order variational formulae given in Section 6 are valid for generalized Hamiltonians.

Although the principal practical utility of generalized variational Hamiltonians is in connection with a partial determination of the optimal control vector (so $\mathbf{v}$ is of smaller dimensionality than $\mathbf{u}$) such Hamiltonians may also arise from the use of generalized control transformations (i.e., transformations in which the new control vector depends on the adjoint vector $\mathbf{p}$, as well as on $\mathbf{x}$, t, and the old control vector) in which the dimensionality of the control vector does not change. Such transformations are evidently of both theoretical and practical interest. Some of the uses of generalized control transformations (with or without reduction of control dimensionality) are illustrated in the following sections. It should perhaps be emphasized again that these control transformations do not affect either the state vector $\mathbf{x}(t)$ or the adjoint vector $\mathbf{p}(t)$. Consequently one may, if desired, apply them to particular subarcs of a trajectory without affecting the treatment of the rest of the trajectory.

Still another way in which generalized variational Hamiltonians may arise is by canonical transformations which alter the state and adjoint variables while preserving equations (5), (8), and (15). Discussion of these more general transformations is outside the scope of the present paper.

4. Singular arcs of linear type

In this section, the use of generalized variational Hamiltonians will be illustrated by a discussion and classification of cases in which the singular control variables are all of the "linear" type. (This classification will subsequently be shown to be relevant for singular arcs in general.) Certain quantities defined in the course of the discussion will be useful in the next section, where the generalized Legendre-Clebsch condition is stated and discussed.

Consider a Hamiltonian $H(\mathbf{x}, \mathbf{p}, \mathbf{u}, t)$ which is regular in the region of interest, i.e., it and all its partial derivatives, up to whatever order is needed, exist and are continuous in this region. Throughout the region of interest, the matrix $\mathbf{H}_{uu}$ is assumed to have rank $n_c - r$. Then by well-known implicit function theorems, the equations $\mathbf{H}_u = 0$ can be used to express all the original control variables as regular functions of $\mathbf{x}$, $\mathbf{p}$, t and of r new control variables (which may possibly be a subset of the original control variables). In other words, the $n_c - r$ nonsingular control variables can be eliminated, leaving only the r singular control variables. Also, the resulting new Hamiltonian will be regular. Assume that the new Hamiltonian is linear in the new, r-component control vector. Then

$$H(\mathbf{x}, \mathbf{p}, \mathbf{u}, t) \equiv H(\mathbf{x}, \mathbf{p}, 0, t) + \mathbf{u}^T\mathbf{H}_u(\mathbf{x}, \mathbf{p}, t). \quad (16)$$

The maximum principle requires $\mathbf{H}_u$ to vanish identically on any extremal arc where $\mathbf{u}$ is in the interior of its allowed region. This does not directly determine $\mathbf{u}$, since $\mathbf{H}_u$ is independent of $\mathbf{u}$ by the linearity assumption. However, (as is well known) it determines $\mathbf{u}$ indirectly, via the time derivatives of $\mathbf{H}_u$. The identical vanishing of $\mathbf{H}_u$, which is necessary to keep the trajectory singular, gives the sequence of conditions

$$\mathbf{H}_u \equiv 0, \qquad (\mathbf{H}_u)^{\cdot} \equiv 0, \qquad (\mathbf{H}_u)^{\cdot\cdot} \equiv 0, \quad \text{etc.} \quad (17)$$

The condition $\mathbf{H}_u \equiv 0$ is independent of $\mathbf{u}$ (as already noted) and hence gives a relation among $\mathbf{x}$, $\mathbf{p}$, and t. By use of the equations $\dot{\mathbf{x}} = \mathbf{H}_p$ and $\dot{\mathbf{p}} = -\mathbf{H}_x$, the other conditions given in (17) can be successively reduced to similar relations among these variables, until sooner or later (in general) a relation will be encountered which explicitly involves $\mathbf{u}$. Let $\mathbf{Q}_m$ denote the $r \times r$ matrix whose elements are

$$Q_{mij} = \frac{\partial}{\partial \mathbf{u}}\left[\left(\frac{d}{dt}\right)^m \frac{\partial H}{\partial \mathbf{u}}\right], \quad (18)$$

and let M denote the smallest value of m for which $\mathbf{Q}_m$ has at least one nonzero element. In general, M is a function of $\mathbf{x}$, $\mathbf{p}$, and t, but to simplify the discussion we shall assume that M is constant in the neighborhood of the extremal arc of interest, and make a similar assumption for the rank of $\mathbf{Q}_M$. These assumptions exclude certain atypical cases in which the extremal arc coincides with a line or surface in the $\mathbf{x}$, $\mathbf{p}$, or t space where M is greater, or the rank of $\mathbf{Q}_M$ is less, than at neighboring points. (I am indebted to the anonymous referee who brought these atypical cases to my attention.) These atypical cases will be discussed in Section 8. In all other cases, M is the first value of m for which the elements of $\mathbf{Q}_m$ do not all vanish identically in the region of interest. For singular arcs, $M \geq 1$ necessarily, and the first of the conditions listed in (17) that explicitly involves $\mathbf{u}$ is

$$0 = (d/dt)^M\mathbf{H}_u = \mathbf{W}_M(\mathbf{x}, \mathbf{p}, t) + \mathbf{Q}_M(\mathbf{x}, \mathbf{p}, t)\mathbf{u}, \quad (19)$$

where $\mathbf{W}_M$ and $\mathbf{Q}_M$ are independent of $\mathbf{u}$ (as indicated) since $\mathbf{u}$ necessarily enters *linearly* into the equation. If $\mathbf{Q}_M$ is nonsingular, the arc is singular of degree M with respect to each of the r singular control variables. (*Note: Degree* of singularity, as defined here, should not be confused with *rank* of singularity, defined earlier as the number of singular control variables.) The optimal control is given by

$$\mathbf{u} = -[\mathbf{Q}_M(\mathbf{x}, \mathbf{p}, t)]^{-1}\mathbf{W}_M(\mathbf{x}, \mathbf{p}, t). \quad (20)$$

The identical vanishing of $(d/dt)^M\,\mathbf{H}_u$ on the arc guarantees the vanishing of all higher derivatives of $\mathbf{H}_u$, so no further relations need be considered. For an arc which is singular of degree M with respect to each of its r singular control variables, (17) therefore gives one relation [Equation (19) or (20)] determining the control vector, and the M relations

$$(d/dt)^m\mathbf{H}_u = \mathbf{W}_m(\mathbf{x}, \mathbf{p}, t) = 0 \quad \text{for} \quad 0 \leq m < M. \quad (21)$$

If M is finite (i.e., if some derivative of $\mathbf{H}_u$ involves $\mathbf{u}$ explicitly) these M relations are necessarily independent, for if any derivative of $\mathbf{H}_u$ below the M^{th} could be expressed as a function of lower-order derivatives (including $\mathbf{H}_u$ itself) with coefficients functions of time, then by successive differentiations the same property could be proved for the M^{th} derivative, and this derivative would therefore be independent of $\mathbf{u}$, contrary to the definition of M. Since the $2n$ components of $\mathbf{x}$ and $\mathbf{p}$ cannot satisfy more than $2n$ independent conditions it follows that either $rM \leq 2n$, or *all* derivatives of $\mathbf{H}_u$ are independent of $\mathbf{u}$. The latter alternative implies an indeterminate control function, a degenerate case that will not be considered.

If the matrix $\mathbf{Q}_M$ is singular, so its rank r_1 is less than r, the arc is singular of degree M with respect to r_1 of its control variables, and singular to some higher degree with respect to other control variables. By use of (19), r_1 of the singular control variables can be expressed as functions of $\mathbf{x}$, $\mathbf{p}$, t and the remaining ("more singular") control variables. Substitution into the Hamiltonian gives a new Hamiltonian with fewer control variables. The new $\mathbf{H}_u$ and its first $M - 1$ derivatives automatically vanish as consequences of relations already obtained. The M^{th} derivative does not involve control variables, but some higher derivative will (in general). This gives a relation like (19) but with a new and larger integer M, and a new matrix $\mathbf{Q}_M$ of smaller size. If the new $\mathbf{Q}_M$ is nonsingular, the process terminates, otherwise it continues in an obvious manner. Let the sequence of values of M found be M_1, M_2, etc. Then the arc is singular of degree M_1 with respect to r_1 control variables, of degree M_2 with respect to r_2 control variables, and so on. The total degree of singularity of the arc is defined to be

$$M_T = \sum_i r_i M_i. \quad (22)$$

Besides the control equations, there are M_T conditions involving $\mathbf{x}$, $\mathbf{p}$, and t only, which must be satisfied on the singular arc. By an extension of an argument given earlier, these relations are necessarily independent (implying $M_T \leq 2n$) except in degenerate cases where at least one control variable can never be determined.

5. The generalized Legendre-Clebsch condition

It is convenient to extend the definition of the matrices $\mathbf{Q}_m$ as given by (22) to cases where control variables may enter nonlinearly into the Hamiltonian, so $\mathbf{Q}_m$ may depend on $\mathbf{u}$ as well as on $\mathbf{x}$, $\mathbf{p}$, and t. Let M denote the smallest integer for which $\mathbf{Q}_M$, evaluated on the extremal of interest and with the optimal control, has at least one nonzero element. Then the classical Legendre-Clebsch condition, which requires the matrix $\mathbf{H}_{uu} = \mathbf{Q}_0$ to be negative semi-definite, can be expressed as a special case of a more general necessary condition for optimality, involving $\mathbf{Q}_M$. This condition, which constitutes the complete generalization of the classical Legendre-Clebsch condition for arcs which may be singular or nonsingular, can be stated in the form of two subconditions:

(1) The integer M must be *even*.
(2) If $M = 2k$, then $(-1)^k \mathbf{Q}_{2k}$ must be negative semi-definite.

These two subconditions will be derived in Section 7, using results from Section 6. Their systematic application to a particular problem proceeds as follows. First form $\mathbf{Q}_0 = \mathbf{H}_{uu}$. Then there are three alternatives:

(A) $\mathbf{Q}_0 \neq 0$ and nonsingular
(B) $\mathbf{Q}_0 \neq 0$ and singular
(C) $\mathbf{Q}_0 = 0$.

If $\mathbf{Q}_0 \neq 0$, then $M = 0$, and the generalized condition reduces to the classical Legendre-Clebsch condition, which is automatically satisfied as a consequence of the maximum principle. If $\mathbf{Q}_0$ is not only nonzero but nonsingular (alternative A) the arc is nonsingular, the Legendre-Clebsch condition is satisfied in its strong form, and no further tests are necessary. If $\mathbf{Q}_0$ is nonzero but singular (alternative B) the arc is singular, and use of the maximum principle to eliminate nonsingular controls (as described in Section 4 gives a new problem with a control vector of reduced dimensionlity, and such that the new $\mathbf{Q}_0$ is zero. If $\mathbf{Q}_0 = 0$, either originally (alternative C) or as a result of the reduction just mentioned, the procedure is to form successive time-derivatives of $\mathbf{H}_u$ until some derivative (the M^{th}) gives a nonzero $\mathbf{Q}_M$. If M is odd, the arc is non-optimal by subcondition (1) and need not be tested further. If M is even, the matrix $\mathbf{Q}_M$ must be examined to see whether it satisfies subcondition (2). If it does not, the arc is non-optimal and need not be tested further. If $\mathbf{Q}_M$ satisfies subcondition (2) and also is nonsingular, the test ends with a favorable result. If $\mathbf{Q}_M$ satisfies subcondition (2) but is singular, further testing is required. In the usual case where the singular control variables are all of the linear type, the procedures described in Section 4 may be applied to eliminate successive sets of control variables. Each elimination gives a new integer M and a new matrix $\mathbf{Q}_M$ (of reduced size) to which subconditions (1) and (2) may be applied again. The process continues until some test is not met, or until subcondition (2) is satisfied in its strong form (i.e., with a nonsingular $\mathbf{Q}_M$). If some or all of the singular control variables are of the nonlinear type, or if one of the atypical cases mentioned in Section 4 arises, a modified procedure must be used. This will be explained in Section 8.

It will be shown in Section 7 that the matrix $\mathbf{Q}_M$ is necessarily symmetric if M is even, and antisymmetric if M is odd. When there is only one control variable (originally, or after reduction) the matrix $\mathbf{Q}_M$ reduces to a scalar q_M, and M is necessarily even since an antisymmetric 1×1 matrix vanishes. The generalized Legendre-Clebsch condition then reduces to

$$(-1)^k q_{2k} \leq 0, \tag{23}$$

which is the one-variable necessary condition quoted in the Introduction.

If this one-variable necessary condition can be proved, then subcondition (2), which is its obvious multivariable generalization, follows immediately. To see this, consider any arc which is singular of order $2k$ and violates subcondition (2). That is, at least one of the eigenvalues of $\mathbf{Q}_{2k}$ has the wrong sign. Because of the symmetry of $\mathbf{Q}_{2k}$ (to be proved in Section 7) it is possible to diagonalize $\mathbf{Q}_{2k}$ by a transformation of the control variables. The coefficients of this transformation may be functions of t alone, since the diagonalization is required to hold only on the extremal being tested. From the new control variables, select one which corresponds to a wrong-sign eigenvalue, and "freeze" all the other control variables by considering them to be given functions of time. This creates a new problem, with only one control variable. The one-variable condition (generalized Legendre condition) is violated, so if it is a necessary condition for optimality, the arc being tested must be non-optimal for the new problem. But this means that it must be non-optimal for the original problem also, since every control variation permitted in the new problem is also permitted in the original problem. Therefore, the one-variable necessary condition implies the necessity of subcondition (2).

From this discussion it is evident that the generalization from a single control variable to multiple control variables adds nothing essentially new, except the possibility of arcs with odd-degree singularity. These will be shown in Section 7 to be non-optimal, as implied by subcondition (1). A general expression for $\mathbf{Q}_1$, valid for singular cases of linear type with nonsingular controls eliminated, is

$$\mathbf{Q}_1 = \mathbf{H}_{up}\mathbf{H}_{xu} - \mathbf{H}_{ux}\mathbf{H}_{pu}. \tag{24}$$

If α and β are two components of $\mathbf{u}$, the optimality condition $\mathbf{Q}_1 = 0$ can be written out in component form as

$$0 = \sum_1^n \left[\frac{\partial^2 H}{\partial\alpha\,\partial \mathbf{p}_k} \frac{\partial^2 H}{\partial\beta\,\partial \mathbf{x}_k} - \frac{\partial^2 H}{\partial\alpha\,\partial \mathbf{x}_k} \frac{\partial^2 H}{\partial\beta\,\partial \mathbf{p}_k} \right]. \tag{25}$$

By this test, a rank-two singular arc arising in the theory of mid-course corrections of space trajectories has been shown to be non-optimal, verifying a conjecture of Breakwell.[9]

6. Variational Formulae

In this section, certain variational formulae will be derived which relate control variations to changes of the final state. It will be recalled that for weak variations, the control-vector change $\delta\mathbf{u}$ is assumed to be small of first order in some parameter ϵ. For strong variations, the absolute integral of $\delta\mathbf{u}$ is assumed to be small of first order in ϵ, but no restrictions (other than those given by the control constraints) are imposed on $\delta\mathbf{u}$ itself. For either strong or weak variations, the changes of the state and adjoint vectors are of first order in ϵ. The change in final state is to be determined to second order in ϵ.

Consider a given trajectory, which will be referred to as the unperturbed trajectory. A second trajectory (the perturbed trajectory) is to be compared with this trajectory. Let quantities evaluated on the perturbed trajectory be identified by asterisks, whereas the corresponding unstarred quantities are understood to be evaluated on the unperturbed trajectory. Thus $\mathbf{x}^*$, $\mathbf{p}^*$, and $\mathbf{u}^*$ are perturbed variables, whereas $\mathbf{x}$, $\mathbf{p}$, and $\mathbf{u}$ are unperturbed. Define

$$\delta\mathbf{x}(t) = \mathbf{x}^*(t) - \mathbf{x}(t) \tag{26}$$

with analogous definition for the other quantities. It will also be convenient to introduce the notation

$$\tilde{\mathbf{x}} = \mathbf{x} + (1/2)\,\delta\mathbf{x} = \mathbf{x}^* - \tfrac{1}{2}\,\delta\mathbf{x}$$

with analogous definitions for $\tilde{\mathbf{p}}$ and $\tilde{\mathbf{u}}$. Then by using (8) and (15), and relations derived from the latter equation by partial differentiation, it is easy to verify that

$$(\tilde{\mathbf{p}}^T\,\delta\mathbf{x})^{\cdot} = H^* - H - \tfrac{1}{2}(\mathbf{H}_x^* + \mathbf{H}_x)^T\,\delta\mathbf{x} - \tfrac{1}{2}(\mathbf{H}_p^* - \mathbf{H}_p)^T\,\delta\mathbf{p} \tag{27}$$

without approximation. For weak variations, expanding the functions on the right side of this equation in Taylor's series about $\tilde{\mathbf{x}}$, $\tilde{\mathbf{p}}$, $\tilde{\mathbf{u}}$, and noting that the quadratic terms cancel, gives

$$(\tilde{\mathbf{p}}^T\,\delta\mathbf{x})^{\cdot} = [\mathbf{H}_u(\tilde{\mathbf{x}}, \tilde{\mathbf{p}}, \tilde{\mathbf{u}}, t)]^T\,\delta\mathbf{u} + O(\epsilon^3) \tag{28}$$

and integrating this from t_0 to t_f gives

$$[\tilde{\mathbf{p}}^T\,\delta\mathbf{x}]_{t_0}^{t_f} = \int_{t_0}^{t_f} [\mathbf{H}_u(\tilde{\mathbf{x}}, \tilde{\mathbf{p}}, \tilde{\mathbf{u}}, t)]^T\,\delta\mathbf{u}\,dt + O(\epsilon^3), \tag{29}$$

which is the desired relation for weak variations. By a tedious but straightforward computation, it can be shown that the analogous formula

$$[\tilde{\mathbf{p}}^T\,\delta\mathbf{x}]_{t_0}^{t_f} = \int_{t_0}^{t_f} [H(\tilde{\mathbf{x}}, \tilde{\mathbf{p}}, \mathbf{u}^*, t) - H(\tilde{\mathbf{x}}, \tilde{\mathbf{p}}, \mathbf{u}, t)]\,dt + O(\epsilon^3) \tag{30}$$

is valid for strong variations. But this formula, though of considerable intrinsic interest, is not required for present purposes. Equation (29) can be expressed in the equivalent form

$$[(\mathbf{p} + \tfrac{1}{2}\,\delta\mathbf{p})^T\,\delta\mathbf{x}]_{t_0}^{t_f} = \int_{t_0}^{t_f} (\mathbf{H}_u + \tfrac{1}{2}\,\delta\mathbf{H}_u)^T\,\delta\mathbf{u}\,dt + O(\epsilon^3). \tag{31}$$

So far, no assumption has been made that the unperturbed trajectory is an extremal. Introducing this assumption gives $\mathbf{H}_u = 0$, so (31) simplifies to

$$[(\mathbf{p} + \tfrac{1}{2}\,\delta\mathbf{p})^T\,\delta\mathbf{x}]_{t_0}^{t_f} = \frac{1}{2}\int_{t_0}^{t_f} (\delta\mathbf{H}_u)^T\,\delta\mathbf{u}\,dt + O(\epsilon^3). \tag{32}$$

For weak variations it is legitimate to assume (as cannot be assumed for strong variations) that $\delta\mathbf{u}$, $\delta\mathbf{x}$, and $\delta\mathbf{p}$ are expressible as power series in the parameter ϵ:

$$\delta\mathbf{u} = \epsilon\,\delta\mathbf{u}_{(1)} + \tfrac{1}{2}\epsilon^2\,\delta\mathbf{u}_{(2)} + \cdots \tag{33a}$$

$$\delta\mathbf{x} = \epsilon\,\delta\mathbf{x}_{(1)} + \tfrac{1}{2}\epsilon^2\,\delta\mathbf{x}_{(2)} + \cdots \tag{33b}$$

$$\delta\mathbf{p} = \epsilon\,\delta\mathbf{p}_{(1)} + \cdots \tag{33c}$$

Using these expansions, both sides of (32) are to be evaluated to order ϵ^2. To this order, the right side of the equation is

$$\tfrac{1}{2}\epsilon^2 \int_{t_0}^{t_f} (\delta\mathbf{H}_{u(1)})^T\,\delta\mathbf{u}_{(1)}\,dt, \tag{34}$$

where

$$\delta\mathbf{H}_{u(1)} = \mathbf{H}_{ux}\,\delta\mathbf{x}_{(1)} + \mathbf{H}_{up}\,\delta\mathbf{p}_{(1)} + \mathbf{H}_{uu}\,\delta\mathbf{u}_{(1)}. \tag{35}$$

The quantities $\delta\mathbf{x}_{(1)}$ and $\delta\mathbf{p}_{(1)}$ are solutions of the following pair of linear differential equations:

$$\delta\dot{\mathbf{x}}_{(1)} = \mathbf{H}_{px}\,\delta\mathbf{x}_{(1)} + \mathbf{H}_{pp}\,\delta\mathbf{p}_{(1)} + \mathbf{H}_{pu}\,\delta\mathbf{u}_{(1)} \tag{36}$$

$$\delta\dot{\mathbf{p}}_{(1)} = -\mathbf{H}_{xx}\,\delta\mathbf{x}_{(1)} - \mathbf{H}_{xp}\,\delta\mathbf{p}_{(1)} - \mathbf{H}_{xu}\,\delta\mathbf{u}_{(1)}. \tag{37}$$

Let us consider solutions of (36) and (37) for which $\delta\mathbf{x}_{(1)}$ vanishes at times t_0 and t_f, so there is no first order state-perturbation at the endpoints. Then to order ϵ^2, (32) can be rewritten as

$$[\mathbf{p}^T\,\delta\mathbf{x}]_{t_0}^{t_f} = \tfrac{1}{2}\epsilon^2 \int_{t_0}^{t_f} (\delta\mathbf{H}_{u(1)})^T\,\delta\mathbf{u}_{(1)}\,dt, \tag{38}$$

where it is understood that $\delta\mathbf{x}$ is $O(\epsilon^2)$ at time t_0 and t_f.

Let us assume for simplicity that the endtimes t_0 and t_f are fixed, and that the state constraints (if any) imposed at times t_0 and t_f are equality constraints. Then if $\mathbf{P}[\mathbf{x}(t_0), \mathbf{x}(t_f)]$ is the quantity to be extremized, the transversality conditions which $\mathbf{p}(t)$ is required to satisfy at the endtimes guarantee the existence of a constant μ(with $\mu \geq 0$ if $\mathbf{P}$ is to be maximized, and $\mu \leq 0$ if $\mathbf{P}$ is to be minimized) such that

$$\mu \, \delta\mathbf{P} = [\mathbf{p}^T \, \delta\mathbf{x} + O(|\delta\mathbf{x}|^2)]_{t_0}^{t_f} \tag{39}$$

for all endpoint perturbations which are consistent with the state constraints referred to above. Putting this into (38), and dropping the terms in $|\delta\mathbf{x}(t_0)|^2$ and $|\delta\mathbf{x}(t_f)|^2$ because they are of order ϵ^4, gives

$$\mu \, \delta\mathbf{P} = \tfrac{1}{2}\epsilon^2 \int_{t_0}^{t_f} (\delta\mathbf{H}_{u(1)})^T \, \delta\mathbf{u}_{(1)} \, dt \tag{40}$$

to order ϵ^2. If the trajectory is a local optimum, $\mu\,\delta\mathbf{P}$ must be ≤ 0 for ϵ sufficiently small. Therefore, a necessary condition for optimality is that the integral on the right side of this equation must be nonpositive for any solution of (35), (36), and (37) with $\delta\mathbf{x}_{(1)}$ vanishing at the endpoints.

This condition is essentially the second-variational condition which is derived in the classical variational calculus in connection with Jacobi's accessory minimum problem. In the next section, it will be used to derive the generalized Legendre-Clebsch condition which is the main result of this paper.

7. Proof of the generalized Legendre-Clebsch condition

In this section, the generalized Legendre-Clebsch condition will be shown to be a necessary condition for optimality, by use of the method of special variations. That is, it will be shown that if the generalized Legendre-Clebsch condition is violated at any time between t_0 and t_f, then the trajectory cannot be optimal because a variation of the state, adjoint, and control variables can be found which satisfies the conditions given in Section 6 and gives a positive value for the integral on the right side of (40). The basic method is due to Kelley[1] but has been improved and generalized.

Before beginning the proof, it is convenient to introduce some abbreviated notation. Let Δ denote the value of the integral in (40), and let $\delta\mathbf{u}_{(1)}$ and $\delta\mathbf{H}_{u(1)}$ be represented by $\mathbf{v}$ and $\mathbf{K}$ respectively. Then

$$\Delta = \int_{t_0}^{t_f} \mathbf{K}^T(t)\mathbf{v}(t) \, dt. \tag{41}$$

Also, let $\mathbf{z}$ be a vector with $2n$ components, of which the first n give $\delta\mathbf{x}_{(1)}$ and the second n give $\delta\mathbf{p}_{(1)}$, and introduce the $2n \times r$ matrix $\mathbf{R}$ defined by

$$\mathbf{R} = \begin{bmatrix} \mathbf{H}_{xu} \\ \mathbf{H}_{pu} \end{bmatrix} \tag{42}$$

and the $2n \times 2n$ matrices $\mathbf{S}$, $\mathbf{J}$, $\mathbf{I}$ defined by

$$\mathbf{S} = \begin{bmatrix} \mathbf{H}_{xx} & \mathbf{H}_{xp} \\ \mathbf{H}_{px} & \mathbf{H}_{pp} \end{bmatrix} = \mathbf{S}^T \tag{43}$$

$$\mathbf{J} = \begin{bmatrix} \mathbf{O}_n & \mathbf{I}_n \\ -\mathbf{I}_n & \mathbf{O}_n \end{bmatrix} = -\mathbf{J}^T \tag{44}$$

$$\mathbf{I} = \begin{bmatrix} \mathbf{I}_n & \mathbf{O}_n \\ \mathbf{O}_n & \mathbf{I}_n \end{bmatrix} = -\mathbf{J}^2, \tag{45}$$

where $\mathbf{I}_n$ is the n-dimensional unit matrix and $\mathbf{O}_n$ is the n-dimensional zero matrix. With this notation, (35), (36), and (37) can be rewritten as follows:

$$\mathbf{K} = \mathbf{R}^T\mathbf{z} + \mathbf{H}_{uu}\mathbf{v} \tag{46}$$

$$\dot{\mathbf{z}} = \mathbf{JSz} + \mathbf{JRv}. \tag{47}$$

Now let t_2 denote a time at which the generalized Legendre-Clebsch condition is violated, and consider an interval of width $2T$ centered on this time. The interval begins at time $t_1 = t_2 - T$ and ends at time $t_3 = t_2 + T$. The time t_2 is assumed to be in the interior of a singular arc, and T is to be chosen so small that the entire interval from t_1 to t_3 is also in this arc. For times before t_1, the variations $\mathbf{z}$ and $\mathbf{v}$ are chosen to be identically zero. For times between t_1 and t_3, $\mathbf{v}(t)$ is chosen to be of the following form which depends on M, the order of singularity of the arc:

$$\mathbf{v}(t) = \mathbf{a}^{(k)}(t) \qquad \text{if} \quad M = 2k \tag{48a}$$

$$\mathbf{v}(t) = \mathbf{a}^{(k+1)}(t) \qquad \text{if} \quad M = 2k + 1 \tag{48b}$$

where the superscripts in parentheses denote orders of differentiation. The vector-function $\mathbf{a}(t)$ and its first k derivatives are required to be continuous between times t_1 and t_3 and to vanish at these two times.

As a result of the control variation between times t_1 and t_3, $\mathbf{z}(t_3)$ will generally be nonzero. For times between t_3 and t_f, the control variation is to be chosen in a manner that ensures that the first n components of $\mathbf{z}$ will vanish at time t_f (i.e., so that $\delta\mathbf{x}_{(1)}$ vanishes at time t_f as required by the conditions given in Section 6). One component of $\delta\mathbf{x}_{(1)}$, the component normal to $\mathbf{p}_f$, *automatically* vanishes at time t_f, since $\mathbf{p}^T \, \delta\mathbf{x}_{(1)}$ is a constant and is initially zero. It is explicitly assumed that it is possible to choose $\mathbf{v}(t)$ between t_3 and t_f to make the other $n - 1$ components vanish also. This controllability (normality) assumption is probably not essential for the validity of the generalized Legendre-Clebsch condition, but it is essential for the derivation given here.

It is important to note that the controllability assumption need only be valid *before* the partial elimination of control variables discussed in Section 3. This is because one may use a generalized Hamiltonian and reduced control vector for time $t \geq t_3$, but retain the original Hamiltonian and control vector for times after t_3.

For times between t_3 and t_f, $\mathbf{v}(t)$ and $\mathbf{K}(t)$ are $O(|\mathbf{z}(t_3)|)$ in magnitude, so (41) can be rewritten as

$$\Delta = \int_{t_1}^{t_3} \mathbf{K}^T(\mathbf{v})\mathbf{v}(t)\,dt + O(|\mathbf{z}(t_3)|^2). \tag{49}$$

Using (48) and integrating by parts k times gives

$$\Delta = (-1)^k \int_{t_1}^{t_3} [\mathbf{K}^{(k)}(t)]^T\mathbf{a}(t)\,dt + O(|\mathbf{z}(t_3)|^2) \tag{50a}$$

if M is even, and

$$\Delta = (-1)^k \int_{t_1}^{t_3} [\mathbf{K}^{(k)}(t)]^T\dot{\mathbf{a}}(t)\,dt + O(|\mathbf{z}(t_3)|^2) \tag{50b}$$

if M is odd. Now let us introduce the scalar quantities α and β, defined as follows:

$$\alpha = \int_{t_1}^{t_3} |\mathbf{a}(t)|\,dt \tag{51}$$

$$\beta = \int_{t_1}^{t_3} |\dot{\mathbf{a}}(t)|\,dt. \tag{52}$$

It will be shown that

$$\mathbf{z}(t_3) = O(\alpha) \tag{53}$$

$$\mathbf{K}^{(k)}(t) = \mathbf{Q}_M(t)\mathbf{a}(t) + O(\alpha). \tag{54}$$

Whence, for even M,

$$\Delta = (-1)^k \int_{t_1}^{t_3} \mathbf{a}^T(t)\mathbf{Q}_M(t)\mathbf{a}(t)\,dt + O(\alpha^2) \tag{55a}$$

and for odd M, using the easily proved relation $\alpha \leq T\beta$,

$$\Delta = (-1)^k \int_{t_1}^{t_3} \mathbf{a}^T(t)\mathbf{Q}_M(t)\dot{\mathbf{a}}(t)\,dt + O(\alpha\beta). \tag{55b}$$

From these relations, the generalized Legendre-Clebsch condition follows immediately. To see this, consider first the case of even M and assume that $(-1)^k\mathbf{Q}_M(t_2)$ has a positive eigenvalue, contrary to the generalized Legendre-Clebsch condition. Then choosing $\mathbf{a}(t)$ parallel to the corresponding eigenvector, and making use of the fact (to be proved later) that $\mathbf{Q}_M$ is symmetric when M is even, gives

$$\Delta \geq \lambda \int_{t_1}^{t_3} |\mathbf{a}(t)|^2\,dt + O(\alpha^2), \tag{56}$$

where λ is a lower bound for the eigenvalue, over the interval from t_1 to t_3. Since the eigenvalue is a continuous function of time and is positive at time t_2, this bound can be chosen positive if T is sufficiently small. Also, Schwarz's inequality gives

$$\alpha^2 \leq 2T \int_{t_1}^{t_3} |\mathbf{a}(t)|^2\,dt, \tag{57}$$

which shows that the second term of (56) is of higher order in T than the first term, so if T is sufficiently small, Δ is necessarily positive. Therefore, the trajectory cannot be optimal. This proves the necessity of the generalized Legendre-Clebsch condition for the case of even M. Now consider the case of M odd. For this case, $\mathbf{Q}_M$ is an antisymmetric matrix, as will be shown shortly By the definition of M, $\mathbf{Q}_M$ must have at least one nonzero element. Without loss of generality, let $\mathbf{Q}_{M12} = -\mathbf{Q}_{M21}$ be such an element, and choose $\mathbf{a}(t)$ to have all its components zero except $a_1(t)$ and $a_2(t)$. Then

$$\Delta = (-1)^k \int_{t_1}^{t_3} (a_1\dot{a}_2 - a_2\dot{a}_1)\mathbf{Q}_{M12}\,dt + O(\alpha\beta). \tag{58}$$

The integral can be made to have either sign, since one can reverse the sign of $a_1(t)$ while leaving $a_2(t)$ unchanged. Also its magnitude can readily be made arbitrarily large compared to $\alpha\beta$. This may be seen as follows: define $\Delta t = (t - t_1)$, and choose $a_1(t)$ and $a_2(t)$ to be of the forms

$$a_1(t) = T^{k+1}g_1(\Delta t/T) \tag{59a}$$

$$a_2(t) = T^{k+1}g_2(\Delta t/T) \tag{59b}$$

so the maximum magnitude of $\mathbf{v} = \mathbf{a}^{(k+1)}$ will be independent of T. Then α varies like T^{k+2} and β varies like T^{k+1} so $\alpha\beta$ varies like T^{2k+3}, whereas the integral varies like $\mathbf{Q}_{M12}(t_2)$ times T^{2k+2}, plus terms of higher order, and hence dominates if T is sufficiently small. Therefore Δ can be made to have either sign, so the trajectory cannot be optimal. This proves the necessity of the generalized Legendre-Clebsch condition for the case of odd M, and hence completes the derivation except for proofs of (53) and (54) and of the symmetry properties of $\mathbf{Q}_M$. These proofs are based on certain explicit expressions for $\mathbf{z}(t)$, Δ, and derivatives of $\mathbf{K}(t)$ up to and including the M^{th}. Only cases with $M \geq 1$ (i.e., singular cases with all nonsingular controls eliminated) need be considered. For such cases, $\mathbf{H}_{uu} = 0$ and (46) reduces to

$$\mathbf{K} = \mathbf{R}^T\mathbf{z}. \tag{60}$$

Let us define a sequence of matrices $\mathbf{R}_i$, of dimensionality $2n \times r$, by the equations

$$\mathbf{R}_0 = \mathbf{R} \tag{61a}$$

$$\mathbf{R}_{i+1} = \dot{\mathbf{R}}_i - \mathbf{SJR}_i \quad (i \geq 0), \tag{61b}$$

or in operator form by the equation

$$\mathbf{R}_i = [\mathbf{I}(d/dt) - \mathbf{SJ}]^i\mathbf{R}. \tag{62}$$

It is easy to verify that

$$(\mathbf{R}_i^T\mathbf{z})^{\cdot} = (\mathbf{R}_{i+1})^T\mathbf{z} + \mathbf{R}_i^T\mathbf{JR}_0\mathbf{v}, \tag{63}$$

whence, in particular,

$$\dot{\mathbf{K}} = (\mathbf{R}_0^T\mathbf{z})^{\cdot} = \mathbf{R}_1^T\mathbf{z} + \mathbf{R}_0^T\mathbf{JR}_0\mathbf{v}. \tag{64}$$

If $\mathbf{R}_0{}^T\mathbf{J}\mathbf{R}_0$ is not a null matrix, then $M = 1$ and $\mathbf{Q}_M = \mathbf{R}_0{}^T\mathbf{J}\mathbf{R}_0$. If $\mathbf{R}_0{}^T\mathbf{J}\mathbf{R}_0$ is a null martix, then $\dot{\mathbf{K}} = \mathbf{R}_1{}^T\mathbf{z}$, and using (63) again gives

$$\ddot{\mathbf{K}} = (\mathbf{R}_1^T\mathbf{z})^{\cdot} = \mathbf{R}_2^T\mathbf{z} + \mathbf{R}_1^T\mathbf{J}\mathbf{R}_0\mathbf{v}, \tag{65}$$

which is analogous to (64) but with different subscripts. If $\mathbf{R}_1{}^T\mathbf{J}\mathbf{R}_0$ is not a null matrix, then $M = 2$ and $\mathbf{Q}_M = \mathbf{R}_1{}^T\mathbf{J}\mathbf{R}_0$. If $\mathbf{R}_1{}^T\mathbf{J}\mathbf{R}_0$ is a null matrix, the process continues in an obvious manner. The general result is that

$$(\mathbf{R}_{M-1})^T\mathbf{J}\mathbf{R}_0 = \mathbf{Q}_M \tag{66}$$

$$\mathbf{K}^{(M)} = \mathbf{R}_M^T\mathbf{z} + \mathbf{Q}_M\mathbf{v} \tag{67}$$

and that, for all $i < M$,

$$(\mathbf{R}_{i-1})^T\mathbf{J}\mathbf{R}_0 = 0 \tag{68}$$

$$\mathbf{K}^{(i)} = \mathbf{R}_i^T\mathbf{z}. \tag{69}$$

To demonstrate the symmetry properties of $\mathbf{Q}_M$ it is convenient to introduce a family of matices $\mathbf{A}_{i,j}$ which are all of dimensionality $r \times r$. The indices i, j do not refer to matrix elements, but identify members of the family. The defining equation is

$$\mathbf{A}_{i,j} = \mathbf{R}_i^T\mathbf{J}\mathbf{R}_j \qquad 0 \le i, j \le M-1. \tag{70}$$

It is evident that $(\mathbf{A}_{i,j})^T = -\mathbf{A}_{j,i}$ and that $\mathbf{Q}_M = \mathbf{A}_{M-1,0}$ so $(\mathbf{Q}_M)^T = -\mathbf{A}_{0,M-1}$. Differentiating (70) and using (61b) gives

$$\dot{\mathbf{A}}_{i,j} = \mathbf{A}_{i,j+1} + \mathbf{A}_{i+1,j}. \tag{71}$$

Now let s denote the least value of $i + j$ for which $\mathbf{A}_{i,j}$ is nonzero for some i. Clearly $s \le M-1$ since $\mathbf{A}_{M-1,0} = \mathbf{Q}_M \ne 0$. Since $\mathbf{A}_{s-i-1,i} \equiv 0$, and its derivative is also zero, (71) gives

$$0 = \mathbf{A}_{s-i,i} + \mathbf{A}_{s-i,i+1}. \tag{72}$$

Whence, by induction,

$$\mathbf{A}_{s-i,i} = (-1)^i\mathbf{A}_{s,0}. \tag{73}$$

The integer s must be equal to $M-1$ since (as already stated) it cannot be greater than $M-1$, and by (66) and (73) it cannot be less. Setting $i = s = M-1$ in (73) gives

$$\mathbf{A}_{0,M-1} = (-1)^{M-1}\mathbf{A}_{M-1,0}, \tag{74}$$

which is equivalent to

$$\mathbf{Q}_M^T = (-1)^M\mathbf{Q}_M. \tag{75}$$

This completes the proof of the symmetry properties of $\mathbf{Q}_M$ for $M \ge 1$. (When $M = 0$, $\mathbf{Q}_M = \mathbf{H}_{uu}$ which is obviously symmetric, so the symmetry property holds for this case also.) The next step is to derive an explicit expression for $\mathbf{z}(t)$. Let $\mathbf{\Phi}(t, \tau)$ denote the transition matrix associated with (47). That is $\mathbf{\Phi}$ is the $2n \times 2n$ matrix which satisfies the following set of equations:

$$\mathbf{\Phi}(t, t) = \mathbf{I} \tag{76}$$

$$(d/dt)\mathbf{\Phi}(t, \tau) = \mathbf{J}\mathbf{S}(t)\mathbf{\Phi}(t, \tau) \tag{77}$$

$$(d/d\tau)\mathbf{\Phi}(t, \tau) = -\mathbf{\Phi}(t, \tau)\mathbf{J}\mathbf{S}(\tau). \tag{78}$$

In terms of $\mathbf{\Phi}$ the solution of (47) is

$$\mathbf{z}(t) = \int_{t_1}^{t} \mathbf{\Phi}(t, \tau)\mathbf{J}\mathbf{R}(\tau)\mathbf{v}(\tau)\, d\tau, \tag{79}$$

since $\mathbf{z}(t_1) = 0$ by choice. Equations (78) and (62) together give

$$(d/d\tau)[\mathbf{\Phi}(t, \tau)\mathbf{J}\mathbf{R}_i(\tau)] = \mathbf{\Phi}(t, \tau)\mathbf{J}\mathbf{R}_{i+1}(\tau). \tag{80}$$

Integrating the right side of (79) by parts k times, and using (73) and (48a) gives

$$\mathbf{z}(t) = -\sum_{i=0}^{k-1} (-1)^{k-i}\mathbf{J}\mathbf{R}_{k-1-i}(t)\mathbf{a}^{(i)}(t) + (-1)^k \int_{t_1}^{t} \mathbf{\Phi}(t, \tau)\mathbf{J}\mathbf{R}_k(\tau)\mathbf{a}(\tau)\, d\tau \tag{81a}$$

for the case of even M. Similarly, for the case of odd M integrating by parts $k + 1$ times and using (73) and (48b gives

$$\mathbf{z}(t) = +\sum_{i=0}^{k} (-1)^{k-i}\mathbf{J}\mathbf{R}_{k-i}(t)\mathbf{a}^{(i)}(t) + (-1)^{k+1} \int_{t_1}^{t} \mathbf{\Phi}(t, \tau)\mathbf{J}\mathbf{R}_{k+1}(\tau)\mathbf{a}(\tau)\, d\tau. \tag{81b}$$

Setting $t = t_3$ in these formulae, and using the properties of $\mathbf{a}(t)$, gives

$$\mathbf{z}(t_3) = (-1)^{k+1} \int_{t_1}^{t_3} \mathbf{\Phi}(t_3, \tau)\mathbf{J}\mathbf{R}_k(\tau)\mathbf{a}(\tau)\, d\tau \tag{82a}$$

if M is even, and

$$\mathbf{z}(t_3) = (-1)^{k+1} \int_{t_1}^{t_3} \mathbf{\Phi}(t_3, \tau)\mathbf{J}\mathbf{R}_{k+1}(\tau)\mathbf{a}(\tau)\, d\tau \tag{82b}$$

if M is odd. From these equations and (51), Eq. (53) follows immediately. To prove (54) use $k < M$, which by (69) implies $\mathbf{K}^{(k)} = \mathbf{R}_k{}^T\mathbf{z}$. The cases of even M ($M = 2k$) and odd M ($M = 2k + 1$) must be treated separately. Multiplying (81a) and (81b) by $\mathbf{R}_k{}^T$ and using (70) gives

$$\mathbf{K}^{(k)}(t) = \sum_{i=0}^{k-1} (-1)^{k-1-i}\mathbf{A}_{k,k-1-i}\mathbf{a}^{(i)}(t) + \mathrm{O}(\alpha) \tag{83a}$$

for even M, and

$$\mathbf{K}^{(k)}(t) = \sum_{i=0}^{k} (-1)^{k-i}\mathbf{A}_{k,k-i}\mathbf{a}^{(i)}(t) + \mathrm{O}(\alpha) \tag{83b}$$

for odd M. In either case, the sum of the subscripts of $\mathbf{A}$ is $M - 1 - i$. If $i \ne 0$, the subscript-sum is less than $\mathbf{M} - 1$ and the $\mathbf{A}$-matrix vanishes (see discussion following (71)).

Therefore, (83a) and (83b) reduce to

$$\mathbf{K}^{(k)} = (-1)^{k-1}\mathbf{A}_{k,k-1}\mathbf{a}(t) + \mathbf{O}(\alpha) \tag{84a}$$

and

$$\mathbf{K}^{(k)}(t) = (-1)^{k}\mathbf{A}_{k,k}\mathbf{a}(t) + \mathbf{O}(\alpha), \tag{84b}$$

respectively. However, by virtue of (73), $(-1)^{k-1}\mathbf{A}_{k,k-1} = \mathbf{Q}_M$ when $M = 2k$, and $(-1)^{k}\mathbf{A}_{k,k} = \mathbf{Q}_M$ when $M = 2k+1$. Therefore, (84a) and (84b) both reduce to

$$\mathbf{K}^{(k)}(t) = \mathbf{Q}_M(t)\mathbf{a}(t) + \mathbf{O}(\alpha)$$

which is identical with (54), the relation to be proved. This completes the proof of the generalized Legendre-Clebsch condition.

8. Singular arcs of nonlinear type

As has already been mentioned, there exist singular arcs where some or all of the singular control variables appear nonlinearly in the Hamiltonian, and can be directly and uniquely determined by the maximum principle. A simple example is given by the problem of minimizing the integral of $\dot{x}_1{}^2 + x_1{}^4$ from zero to infinity, for a specified value of $x_1(0)$. Applying the usual formalism gives the Hamiltonian

$$H = p_1 u + p_2(x_1^2 + u^4).$$

The absolutely optimal solution consists of a parabolic arc followed by a singular arc where $x_1 \equiv p_1 \equiv u \equiv 0$. This arc is of the nonlinear type.

At first thought, it may seem that singular arcs of nonlinear type must have an artifical and "accidental" character: Since the optimal control is uniquely determined by the maximum principle, there is no control freedom which can be exploited to maintain the singularity condition, so this type of singular arc cannot endure for a finite time unless an identical agreement fortuitously exists between the unique optimal control given by the maximum principle and the control needed to maintain the singularity property.

Further consideration shows that this objection is not necessarily valid. Even if nonlinear singular arcs can only occur for special choices of the state equations, these special cases may be of practical interest. Also, the identical agreement necessary for continuation of the arc may not be fortuitous, but the result of deep identities. This possibility is illustrated by the work of Breakwell[10] who has considered the singular extremals of rocket trajectories in vacuum. In the usual treatment with time as independent variable, the rocket problem has singular extremals of linear type. Breakwell showed that the problem may be transformed so the integral of thrust acceleration becomes the independent variable, and time becomes a control variable. In the transformed problem, the singular extremals are of nonlinear type.

The generalized Legendre-Clebsch condition is as applicable to singular arcs of nonlinear type as it is to other kinds of arc. For example, in the problem just described we find $\mathbf{Q}_2 = 2\mathbf{p}_2$ which satisfies the condition, since $\mathbf{p}_2$ must be chosen negative (otherwise, no meaningful extremals exist). However, in complex cases (which will seldom be encountered in practice) a special difficulty may arise. The derivation of the generalized Legendre-Clebsch condition assumed that the first and second partial derivatives of the Hamiltonian (with respect to components of $\mathbf{x}$, $\mathbf{p}$, and $\mathbf{u}$) exist on the arc to be tested, and are continuous in its neighborhood. If this regularity property holds for the original Hamiltonian, it will hold for the new Hamiltonian generated by the elimination of nonsingular control variables. However, it will generally not continue to hold if the maximum principle is used to eliminate a singular control variable of nonlinear type. This is illustrated by the example just given, where elimination of $\mathbf{u}$ gives

$$H = p_2[x_1^2 - 3(p_1/4p_2)^{4/3}].$$

For this Hamiltonian, the second partial derivative with respect to $\mathbf{p}_1$ does not exist on the singular arc where $p_1 \equiv 0$, so the regularity property has been lost. This does no harm in the example case, since no more control variables remain in the Hamiltonian, so this Hamiltonian need not be used in any "generalized Legendre-Clebsch" tests. But if there are two or more singular control variables, not all occuring linearly in the Hamiltonian, and if the extremal is singular to different degrees with respect to different control variables, a difficulty may arise.

This difficulty can always be resolved by the following procedure. Let $H(\mathbf{x}, \mathbf{p}, \mathbf{u}, t)$ be the Hamiltonian after elimination of all nonsingular control variables. Find a new Hamiltonian which is linear in $\mathbf{u}$, and is such that on the extremal to be tested, the two Hamiltonians give identically the same values for the vectors $\mathbf{H}_x$, $\mathbf{H}_p$, and $\mathbf{H}_u$ and for the matrices $\mathbf{H}_{xx}$, $\mathbf{H}_{xp}$, $\mathbf{H}_{pp}$, $\mathbf{HH}_{pu}$ and $\mathbf{H}_{uu}$. This is always possible: one may, for example, choose the linear Hamiltonian to be

$$H(\mathbf{x}, \mathbf{p}, \mathbf{u}_0, t) + (\mathbf{u} - \mathbf{u}_0)^T\mathbf{H}_u(\mathbf{x}, \mathbf{p}, \mathbf{u}_0, t),$$

where $\mathbf{u}_0(t)$ is the control vector associated with the extremal to be tested. A review of Sections 6 and 7 will show that either Hamiltonian must give the same answer when used to evaluate the second-order effect of an arbitrary, weak control variation about the chosen extremal. Therefore, it is justifiable to use the linear Hamiltonian in place of the original Hamiltonian in testing for optimality. But with the linear Hamiltonian, the procedures discussed in Section 4 become available, making it possible to eliminate sets of control variables (without losing the regularity property) and to apply the entire test sequence described in Section 5.

A slightly more elaborate procedure permits a more uniform treatment of all cases, including the atypical cases mentioned in Section 4. The procedure begins by linearizing about the extremal to be tested, obtaining (35), (36), and (37). The coefficient matrices may be regarded as given functions of time. If all nonsingular control variables have been eliminated, $\mathbf{H}_{uu}$ will be zero. Successive derivatives of $\delta\mathbf{H}_{u(1)}$ can be computed until the newest derivative either fails to be linearly independent of derivatives already found (in which case there is no reason to proceed further, since no additional relations can be found) or involves $\delta\mathbf{u}_{(1)}$ with a nonvanishing coefficient $\mathbf{Q}_M$. After applying the optimality test to $\mathbf{Q}_M$, a number of control variables equal to the rank of $\mathbf{Q}_M$ can be eliminated, and the process continued (if necessary) by computing still higher derivatives. It is easy to verify that the elimination can be done in a way which preserves the forms of (35), (36), and (37). A byproduct of the procedure is a set of independent linear relations among the components of $\delta\mathbf{x}_{(1)}$ and $\delta\mathbf{p}_{(1)}$. Since these relations cannot exceed $2n$ in number, the integer M_T defined in Section 4 must be $\leq 2n$, or the process will end with some components of $\delta\mathbf{u}_{(1)}$ still undetermined. In the latter case, $M_T = \infty$ by convention, and variations higher than the second must be considered to determine optimality.

References

1. H. J. Kelley, "A second variation test for singular extremals," *AIAA Journal* **2**, 1380–1382 (1964).
2. R. E. Kopp and A. G. Moyer, "Necessary conditions for singular extremals," *AIAA Journal* **3**, 1439–1444 (1965).
3. H. J. Kelley, R. E. Kopp and A. G. Moyer, "Singular extremals," in *Optimization—Theory and Applications*, (G. Leitmann, ed), Academic Press, 1966, Vol. 1, Chapter 3.
4. H. M. Robbins, "Optimality of intermediate-thrust arcs of rocket trajectories," *AIAA Journal* **3**, 1094–1098 (1965).
5. A. Miele, "Flight mechanics and variational problems of linear type," *Journal of Aerospace Sciences* **25**, 581–590 (1958).
6. H. J. Kelley, "A transformation approach to singular subarcs in optimal trajectory and control problems," *SIAM Journal on Control* **2**, 234–240 (1964).
7. B. S. Goh, "The second variation for the singular Bolza problem," *SIAM Journal on Control* **2**, 309–325 (1966).
8. C. D. Johnson, "Singular solutions in optimal control problems, in *Advances in Control Systems—Theory and Applications* (C. Leondes, ed.) Academic Press, 1965, Vol. 2.
9. J. V. Breakwell, "A doubly singular problem in optimal interplanetary guidance," *SIAM Journal on Control* **3**, 71–77 (1965).
10. J. V. Breakwell, "Minimum impulse transfer," *AIAA Preprint* 66–416, August 1963; also in *AIAA Progress in Astronautics and Aeronautics: Celestial Mechanics and Astrodynamics*, (V. Szebehely, ed.) Academic Press, New York, 1964, Vol. 14, pp. 583–589.

Received September 5, 1966