SOME RECENT RESULTS IN AEROSPACE VEHICLE TRAJECTORY OPTIMIZATION TECHNIQUES

by

Dr. Edward J. Bauman, Major, USAF
Department of Astronautics, USAF Academy
Colorado Springs, Colorado

Dr. C. T. Leondes, Professor
Department of Engineering
University of California, Los Angeles

Dr. R. A. Niemann, Assistant Professor
Department of Engineering
Los Angeles State College, Los Angeles

Dr. Garrett Paine, Member of the Technical Staff
Jet Propulsion Laboratory
Pasadena, California

A paper presented at the XVIII International
Astronautics Federation Congress,
Belgrade, Yugoslavia, September 1967.

*The research reported herein was supported in part under AFOSR Grant No. AF-AFOSR-669-67 Air Force Office of Scientific Research and NASA Grant 237-62.

TABLE OF CONTENTS

1. INTRODUCTION

Fundamental necessary and sufficient conditions in the calculus of variations, basic to the aerospace vehicle trajectory optimization problem, have been under investigation by mathematicians for many decades, in fact, for hundreds of years. Algorithms for the solution of the optimization problem have been under investigation for many years, but it is only in the last five or ten years that this extremely important area has received intensive effort. There are several fundamental approaches to algorithms for the solution of the aerospace vehicle trajectory optimization problem and this paper presents some original recent results in four areas as studied by the co-authors during the course of their research.

First of all, the computation of optimal trajectories by the method of quasilinearization is presented and it is shown how bounds may be placed on continuous controls directly in the method of quasilinearization without additional variables. Our studies and the studies of individuals in other research laboratories have indicated rather clearly that this method of quasilinearization is one of the most promising methods available for aerospace vehicle trajectory optimization, and therefore the results in this section take on an additional significance as a result. As an example of the techniques developed here the trajectory producing minimum heating of an entry body has been found to demonstrate that the method is generally applicable.

One aspect of aerospace vehicle trajectory optimization emerging with more advanced space systems is that of complex interacting trajectories. The next section presents contributions in this area and applies a multi-level (two-level) technique to optimize interacting trajectories (subarcs) that may contain discontinuities in the state, the state derivatives and/or the cost function at the subarc interfaces. The procedure is to optimize each subarc independently and then to use a 2nd-level controller to adjust the interface conditions to achieve an optimal trajectory. This is an iterative procedure which in this paper uses a 2nd-level gradient controller.

In problems of optimal trajectories the final time T may be fixed or it may be variable. For each fixed time, an optimal trajectory may be determined which minimizes the given performance criterion. The variable time problem involves finding that final time for which the minimized performance criterion is minimal over all other final times. This is the problem considered in the next section. The criteria used to determine the optimum in such techniques as Pontryagin's maximum principle are merely necessary conditions rather than sufficient conditions. In the variable time problem the additional necessary condition that the Hamiltonian $H = 0$ may be satisfied for many different final times as will be demonstrated in this section of the paper. In fact an example will be given in which several local minima and maxima of the cost functional exist. In order to solve the variable time problem, additional techniques are needed to vary the first time T so that the Hamiltonian is driven to zero. Such techniques will be developed in this section.

The paper concludes with a section presenting some new results, techniques, and algorithms for the variable end point problem. In the previous section we considered the problem for which the terminal condition was some fixed constant vector in n-dimensional Euclidean space. This final section considers problems in which the terminal vector is merely constrained to lie in some (n-k) dimensional manifold of n-dimensional Euclidean space.

2. THE COMPUTATION OF OPTIMAL TRAJECTORIES BY THE METHOD OF QUASILINEARIZATION

2.1 Introduction

This section is concerned with the computation of optimal trajectories and shows how bounds may be placed on continuous control directly in the method of quasilinearization without additional variables. The trajectory producing minimum heating of an entry body has been found to demonstrate that the method is generally applicable.

Previous application of the method of quasilinearization to the optimal trajectory problem (see References 2.1, 2.2, 2.3 and others) have used the technique of eliminating the control from the problem by inverting the partial of the pseudo-Hamiltonian with respect to control. In many instances this cannot be done or is inconvenient and results in the addition of extra variables if bounds are placed on the control.[2.4] This difficulty has been cited in Reference 2.5.

The problem to be solved is that of minimizing an integral while satisfying n nonlinear differential equations and boundary conditions. In a Pontryagin Maximum Principle Formulation:

$$
\begin{aligned}
\dot{\underline{x}} &= \underline{f}_1(x,u,t) \quad ; \quad 0 \le t \le T \\
\dot{\underline{\psi}} &= -\frac{\partial H}{\partial x} \\
H &= \underline{\psi}^T \cdot \underline{f}_1 - f_o \\
J &= \int_o^T f_o(x,u)dt \\
u &= u\left\{\min_u H\right\}
\end{aligned}
\tag{2.1}
$$

For notational simplicity the short hand

$$
\underline{y}^T = (x,\psi)^T \tag{2.2}
$$

is employed where y is a 2n vector. Thus

$$\begin{aligned} \dot{\underline{y}} &= \underline{f}(y,u) \\ H &= H(y,u) \\ u &= u\left\{\min_{u} H\right\} \end{aligned} \tag{2.3}$$

It is assumed without loss of generality that the 2n boundary conditions are divided equally between $\underline{y}(0)$ and $\underline{y}(T)$.

2.2 The Modified Method of Quasilinearization is formulated as follows for the n^{th} iteration:

$$\begin{aligned} \dot{\underline{z}}^{n+1} &= \underline{f}\Big|^{n} + \frac{\partial f}{\partial y}\Big|^{n} (\underline{z}^{n+1} - \underline{y}^{n}) + \left(\frac{\partial f}{\partial u}\Big|^{n} \cdot \frac{\partial u}{\partial y}\Big|^{n}\right)(\underline{z}^{n+1} - \underline{y}^{n});\ \underline{z}^{n+1}(0) = \underline{y}^{n}(0) \\ \dot{Z}^{n+1} &= \left(\frac{\partial f}{\partial y}\Big|^{n} + \frac{\partial f}{\partial u}\Big|^{n} \cdot \frac{\partial u}{\partial y}\Big|^{n}\right) Z^{n+1};\ Z^{n+1}(0) = I \\ u^{n} &= u\left\{\min_{u} H\Big|^{n}\right\} \\ \underline{y}^{n+1} &= \underline{z}^{n+1} + Z^{n+1}\,\underline{\alpha}^{n+1} \end{aligned} \tag{2.4}$$

where $\underline{z}$ is a dummy 2n vector, the superscripts indicate iteration number, $g\big|^{n}$ indicates that the function g is evaluated on the n^{th} trajectory, Z is a 2n by 2n matrix, and $\underline{\alpha}$ is a 2n vector.

The constant vector $\underline{\alpha}^{n+1}$ is chosen such that:

$$\left.\begin{aligned} \underline{y}^{n+1}(0) &= \underline{y}^{o}(0) \\ \underline{y}^{n+1}(T) &= \underline{y}^{o}(T) \end{aligned}\right\} \quad \text{where the equality holds only for the components specified by the 2n boundary conditions} \tag{2.5}$$

Only n components of the 2n for $\underline{\alpha}^{n+1}$ need be computed. At $t = 0$

$$\underline{y}^{n+1}(0) = \underline{y}^{n}(0) + \underline{\alpha}^{n+1} \tag{2.6}$$

There are n boundary conditions specified at t=0 which must be satisfied by both $\underline{y}^{n}(0)$ and $\underline{y}^{n+1}(0)$ with the consequence that n components of $\underline{\alpha}^{n+1}$ are zero.

Similarly an nxn matrix formed from the 2n x 2n transition matrix need be inverted in the determination of the n non-zero components of $\underline{\alpha}^{n+1}$:

$$\underline{\alpha}^{*n+1} = \left(Z(T)^{*n+1} \right)^{-1} \left(\underline{y}^{*n+1}(T) - \underline{z}^{*n+1}(T) \right) \tag{2.7}$$

where the star indicates the deletion of the unnecessary components.

If the final time is free than a change of independent variable from t to $\beta\tau$ where β is an undetermined constant and $0 \leq \tau \leq 1$ is used. The constant β is determined in the method of quasilinearization by employing H = 0.

It can be shown theoretically,[2.6] using the Contraction Mapping Principle, that the method converges quadratically under some rather restrictive conditions. To show that the method is practical the minimum heating entry problem was solved for both bounded and unbounded control.

2.3 The Entry Problem

The entry problem formulated is that of finding the control to minimize the convective heating for a low lift to drag entry vehicle. The control is the vehicle attitude.

The state equations are:

$$\begin{aligned}
\dot{v} &= -\frac{S}{2m} \rho v^2 C_D(u) - \frac{g_o \sin\gamma}{(1+\xi)^2} \\
\dot{\gamma} &= \frac{S}{2m} \rho v C_L(u) + \frac{v \cos\gamma}{R(1+\xi)} - \frac{g_o \cos\gamma}{v(1+\xi)^2} \\
\dot{\xi} &= \frac{v}{R} \sin\gamma \\
\dot{\zeta} &= \frac{v}{1+\xi} \cos\gamma \\
\rho &= \rho_o e^{-\beta R \xi} \\
a_p &= \frac{S \rho v^2}{2m\, g_o} \sqrt{C_L^2(u) + C_D^2(u)} \\
|u| &\leq u_1
\end{aligned} \tag{2.8}$$

$$C_D(u) = 1.174 - 0.9 \cos 2u$$

$$C_L(u) = 0.6 \sin 2u \qquad (2.8)\ (\text{Cont.})$$

The boundary conditions are

$$\begin{aligned} v(0) &= 35,000 \text{ ft/sec.} & v(T) &= 27,000 \text{ ft/sec} \\ \gamma(0) &= 8.1^{o} & \gamma(T) &= 0.0^{o} \\ \xi(0) &= 0.0191 & \xi(T) &= 0.0120 \end{aligned} \qquad (2.9)$$

The criterion function is the integral of the convective heating

$$J = \int_{o}^{T} C\, V^{3} \sqrt{\frac{\rho}{N}}\ dt \qquad (2.10)$$

Figure 2.1 illustrates the problem geometry while Table 2.1 defines the nomenclature and the problem constants.

Since the final time is free, the integrations are performed with respect to a dummy variable τ where:

$$t = \beta \tau \quad ; \quad 0 \leq \tau \leq 1 \qquad (2.11)$$

The constant β is determined in the usual manner, employing the additional constraint $H \equiv 0$.

In employing the computational procedure of Equations (2.1) to (2.7) the term $\frac{\partial u}{\partial y}\Big|^{n}$ appears. The term can be found from the partial of the pseudo Hamiltonian with respect to control either directly if the partial can be inverted for $u = u(y,t)$ or by perturbations if it cannot. Both methods have been tried with success in the bounded and unbounded control cases. If the partial cannot be taken then the control u and its perturbations must be found by some other technique, such as the gradient method for functions.

Once a convergent initial guess is found the method converges quadratically until the effects of round-off and truncation errors dominate. The task of obtaining a convergent initial guess for both the state and the adjoint variables can be difficult.

The problem can be alleviated by the simple expedient of employing $\bar{\underline{y}}^{n+1}$ as the $n+1^{st}$ trajectory in place of $\underline{y}^{n+1}$ where

$$\bar{\underline{y}}^{n+1} = (1-\delta)\,\bar{\underline{y}}^{n} + \delta\,\underline{y}^{n+1}, \quad 0 < \delta < 1 \tag{2.12}$$

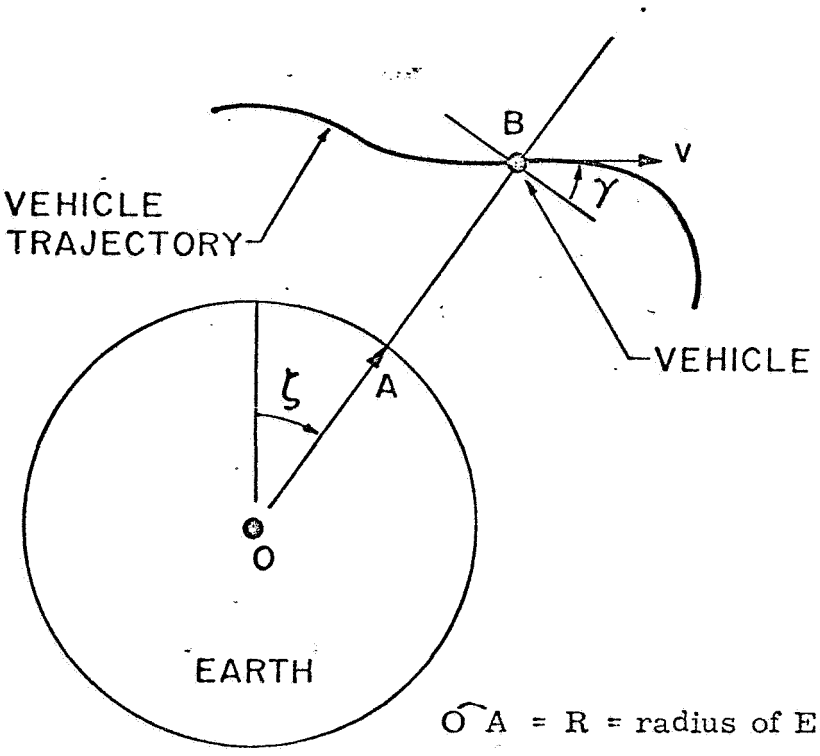
and, where $\underline{y}^{n+1}$ is derived from Equation (2.4) for the first few iterations. If δ is zero then the solution is stationary, if δ is one the usual method of quasilinearization is obtained. For δ's in between convergence is not quadratic, but can be achieved for a wider range of initial trajectories. The increased range of convergence can be shown theoretically.

The solution for the entry problem posed is given in Figures 2.2 and 2.4. In Figure 2.2 the unbounded control is compared to control bounded at $\pm 22.5^{o}$. In Figure 2.3 the time history of the state variables v, γ, ξ is shown. The time history of the corresponding adjoint equations is shown in Figure 2.4. The convergence rates are given in Figure 2.5. It can be seen that the bounded control case converges at the same rate as the unbounded case. The final value of J was 27669 BTU/ft^2 for unbounded control and 32330 BTU/ft^2 for the bounded case.

TABLE 2.1

TABLE OF SYMBOLS IN REENTRY PROBLEMS

v	-	velocity
ξ	-	normalized altitude
γ	-	flight path angle
ζ	-	down range distance
$\frac{S}{M}$	-	ratio of frontal area to vehicle mass $= 0.532\ ft^2/lb$
g_o	-	gravitational constant $= 32.172\ ft/sec^2$
ρ_o	-	air density at sea level $= 0.2704\ d\ 10^{-2}\ slug/ft^3$
β	-	exponential constant $= 0.426 \times 10^{-4} \cdot /ft$
R	-	earth radius $= 20.9 \times 10^6\ ft$
$C_D(u)$	-	drag coefficient
$C_L(u)$	-	lift coefficient
C_{DL}, C_{DO}, C_{LO}	-	drag and lift coefficients
C	-	convective constant $= 0.2 \times 10^{-7}$
N	-	radius of vehicle nose $= 4.0\ ft$
a_p	-	sensed acceleration
u_1	-	control constraint



$\overset{\frown}{OA}$ = R = radius of Earth

$\overset{\frown}{AB}$ = h = altitude of vehicle above Earth

v = velocity vector

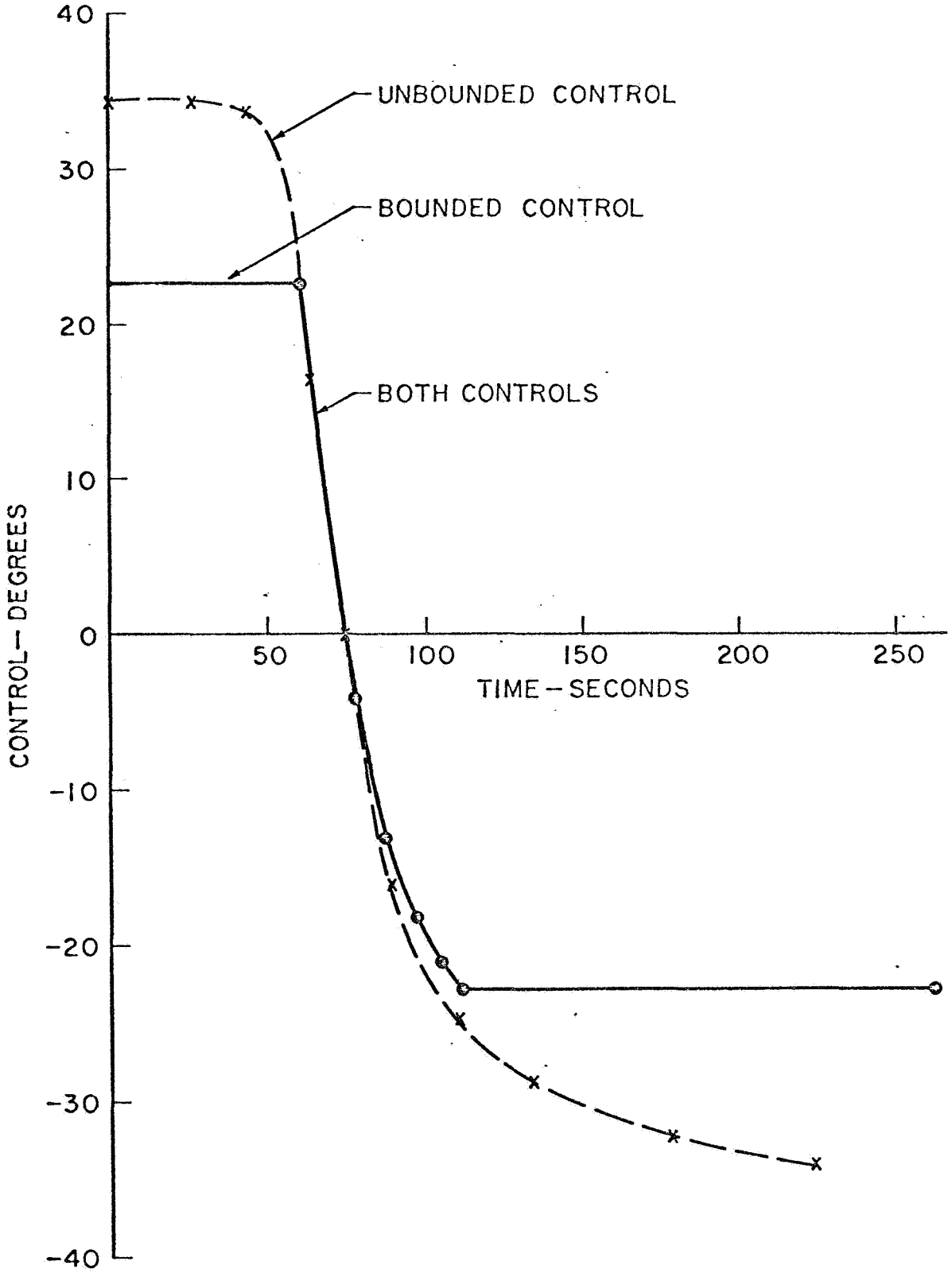
γ = angle between local horizon and velocity vector

ζ = down range distance

$\xi = \frac{h}{R}$ = normalized altitude

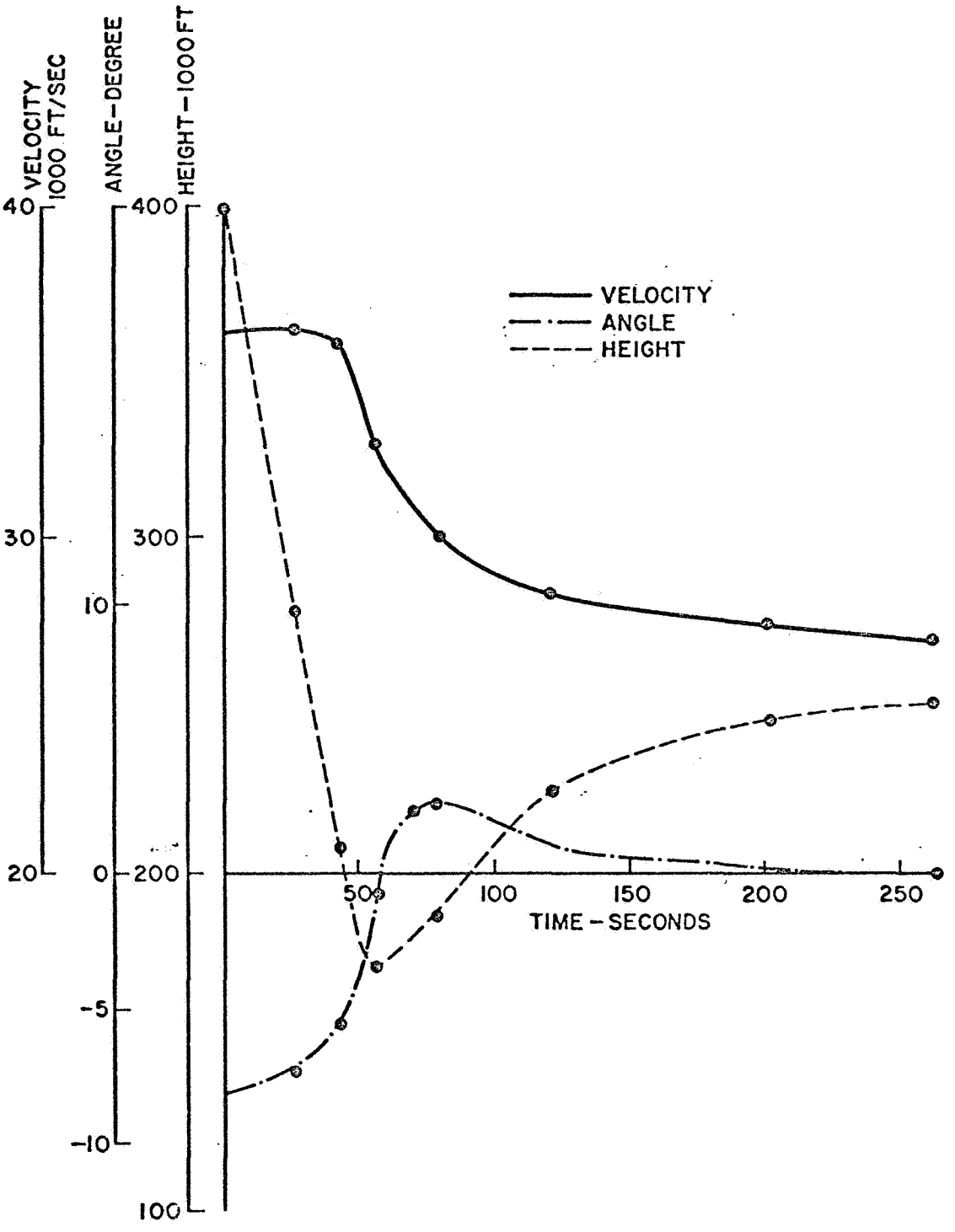
REENTRY VEHICLE COORDINATE SYSTEM

FIGURE 2.1



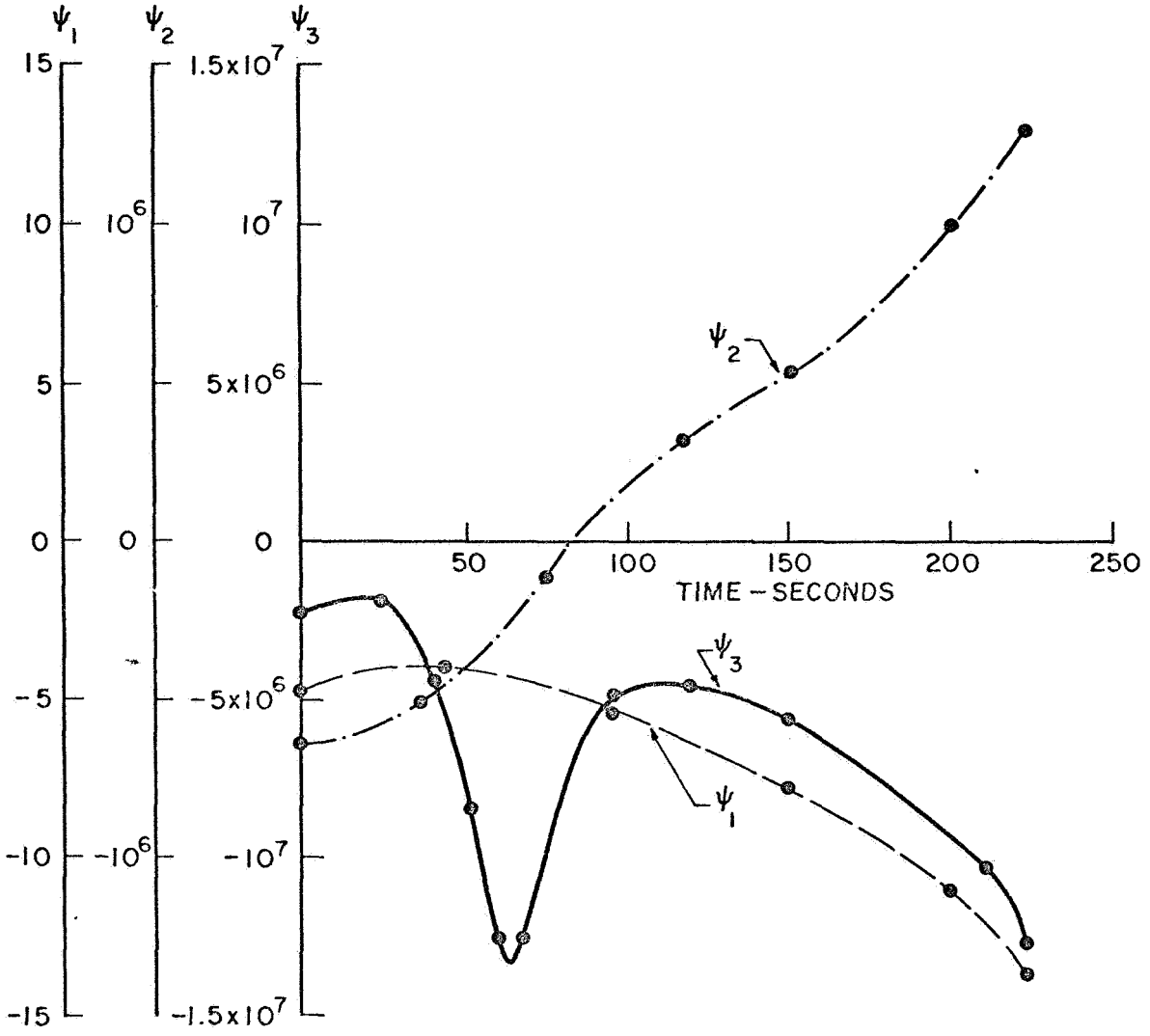
CONTROL HISTORY UNBOUNDED CONTROL
VS. 22.5° BOUND

FIGURE 2.2



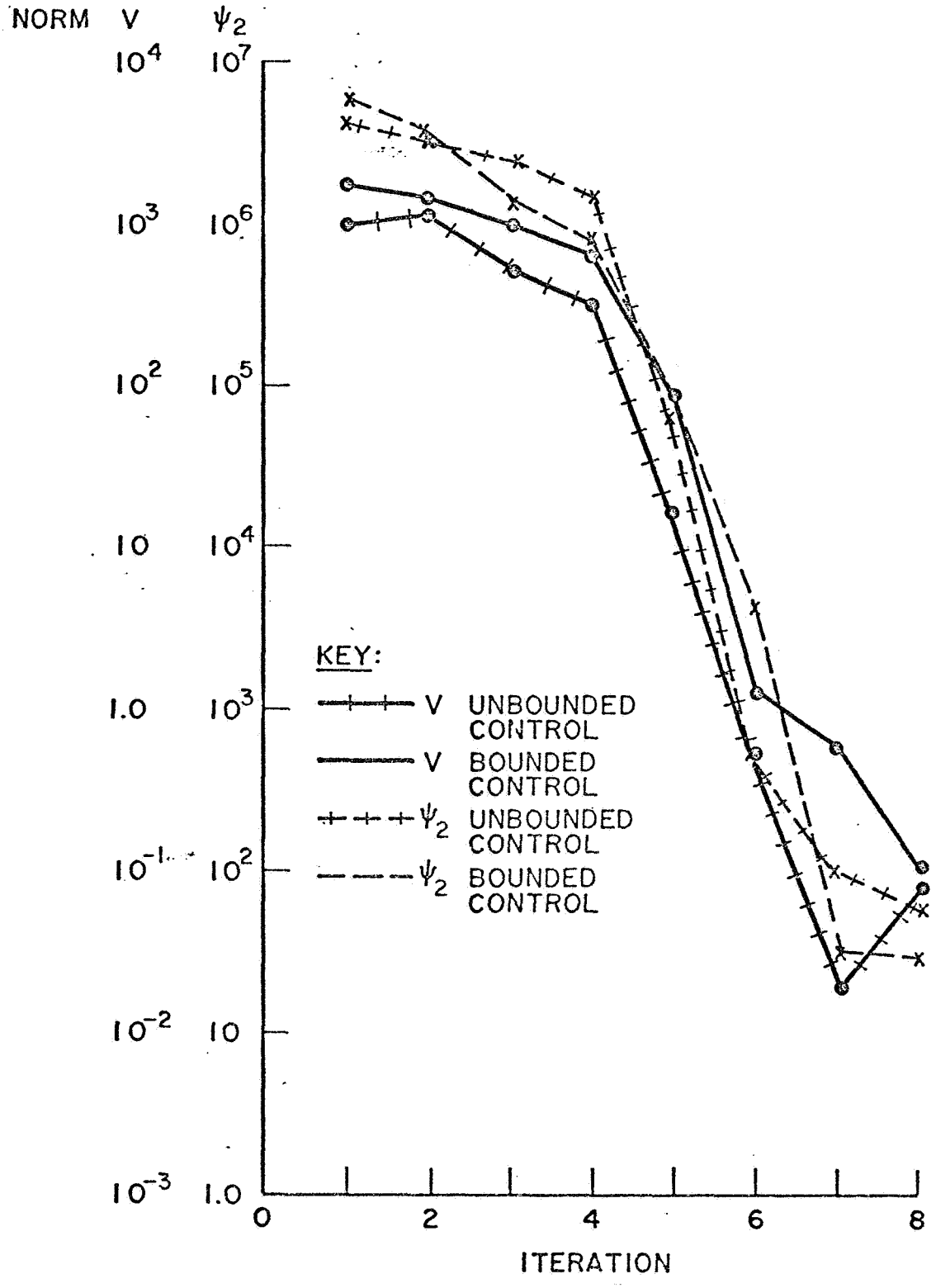
REENTRY VEHICLES STATE VARIABLES
VS. TIME BOUND CONTROL

FIGURE 2.3



REENTRY VEHICLES ADJOINT VARIABLES VS.
TIME BOUNDED CONTROL

FIGURE 2.4



CONVERGENCE RATE FOR BOUNDED CONTROL AND UNBOUNDED CONTROL

FIGURE 2.5

3. MULTI-LEVEL TECHNIQUES APPLIED TO COMPLEX INTERACTING TRAJECTORY SYNTHESIS

3.1 Introduction

This section applies a multi-level (two level) technique to optimize interacting trajectories (subarcs) that may contain discontinuities in the state, the state derivatives and/or the cost function at the subarc interfaces. The procedure is to optimize each subarc independently and then to use a 2nd-level controller to adjust the interface conditions to achieve an optimal trajectory. This is an iterative procedure which in this paper uses a 2nd-level gradient controller.

3.2 Problem Statement

Macko[3.1] extended the multi-level concepts of Lasdon[3.2] and Brosilow[3.3] to decomposition and optimization of non-linear dynamic systems. The results can also be applied to interacting trajectories.[3.4] Consider the optimal control problem of minimizing I with respect to $\underline{u}^i$.

$$I = \sum_{i=1}^{N} \int_{t_o^i}^{t_f^i} F^i(\underline{x}^i, \underline{u}^i, t)\, dt \tag{3.1}$$

subject to

$$\dot{\underline{x}}^i = \underline{f}^i(\underline{x}^i, \underline{u}^i, t^i), \qquad i = 1, \ldots, N \tag{3.2}$$

$$\underline{x}^1(t_o^1) = \underline{x}_o^1 \tag{3.3}$$

$$\underline{\phi}^N(\underline{x}^N, t_f^N) = 0 \tag{3.4}$$

$$\Phi^i = \underline{\phi}^i(\underline{x}^i, t_f^i) + \phi^{i+1}(\underline{x}^{i+1}, t_o^{i+1}) = 0, \tag{3.5}$$

$$i = 1, 2, \ldots, N-1$$

The superscript refers to the subarc number. Each function is assumed continuous and differentiable over its subarc, but discontinuities may occur between subarcs. The interface constraints for the state variables

and times are given by Equation (3.5). If $\underline{x}$ is an n-vector then the $\underline{\psi}$ in Equation (3.5) are n+1-vectors (an interface constraint equation for each state and for time). The constraint equations, (3.2), (3.4) and (3.5) can be added to Equation (3.1) with Lagrange multipliers to form

$$J = I + \sum_{i=1}^{N} \int_{t_o^i}^{t_f^i} (\underline{\psi}^i)^T (\underline{f}^i - \dot{\underline{x}}^i) dt + \sum_{i=1}^{N-1} (\underline{\rho}^i)^T \underline{\Phi}^i + \underline{\nu}^T \underline{\phi}^N \tag{3.6}$$

3.3 Solution Using Multi-Level Optimization

Now J must be minimized with respect to $\underline{u}^i$. By applying the feasible method of decomposition[3.3, 3.4] the initial conditions on each subarc are chosen as known quantities and then J easily decomposes into N separate minimization problems

$$J^i = \int_{t_o^i}^{t_f^i} [F^i + (\underline{\psi}^i)^T (\underline{f}^i - \dot{\underline{x}}^i)] \, dt + (\underline{\rho}^i)^T \underline{\Phi}^i, \quad i = 1, 2, \ldots N-1 \tag{3.7}$$

$$J^N = \int_{t_o^N}^{t_f^N} [F^N + (\underline{\psi}^N)^T (\underline{f}^N - \dot{\underline{x}}^N)] \, dt + \underline{\nu}^T \underline{\phi}^N \tag{3.8}$$

where

$$J = \sum_{i=1}^{N} J^i$$

By satisfying the necessary condition that the first variation of each J^i vanish, all the terms of the first variation of J also vanish with the possible exception of those below:

$$\delta J = \left[(\underline{\psi}^{i+1})^T - (\underline{\rho}^i)^T \frac{\partial \underline{\Phi}^i}{\partial \underline{x}^{i+1}} \right] \underline{\delta x}_o^{i+1}$$

$$+ \left[-F^{i+1} - (\underline{\psi}^{i+1})^T \underline{f}^{i+1} - (\underline{\rho}^i)^T \frac{\partial \underline{\Phi}^i}{\partial t^{i+1}} \right] dt_o^{i+1} \tag{3.9}$$

These are the optimal conditions that may not be satisfied in general because x_o^{i+1} and t_o^{i+1} were fixed for each subarc.

Now since all the constraints have been satisfied by the subarc minimization

$$J = I \tag{3.10}$$

Also since we are minimizing I subject to the constraints, δI should be negative. Therefore, by Equation (3.10) δJ should be made negative. A sufficient condition for $\delta J < 0$ follows from Equation (3.9) if

$$\underline{\delta x}_o^{i+1} = - k \left[(\psi^{i+1})^T - (\underline{\rho}^i)^T - (\underline{\rho}^i)^T \frac{\partial \underline{\Phi}^i}{\partial x^{i+1}} \right] \tag{3.11}$$

$$dt_o^{i+1} = - k \left[-F^{i+1} - (\psi^{i+1})^T \underline{f}^{i+1} - (\underline{\rho}^i)^T \frac{\partial \underline{\Phi}^i}{\partial t^{i+1}} \right] \tag{3.12}$$

Where $k > 0$. The other terms come from the appropriate subarc after it has been minimized. Using Equations (3.11-3.12), for the next iteration let

$$x_o^{i+1}\Big|_{new} = \underline{x}_o^{i+1}\Big|_{old} + \delta \underline{x}_o^{i+1} \tag{3.13}$$

$$t_o^{i+1}\Big|_{new} = t_o^{i+1}\Big|_{old} + dt_o^{i+1} \tag{3.14}$$

Equations (3.13-3.14) give a simple 2nd-level gradient controller that will converge under the stated assumptions toward the minimum. However, in the vicinity of the minimum, k must decrease continuously or this iterative solution method becomes unstable as is true of simple gradient methods.

3.4 An Example

This technique was used to maximize the range of a constant thrust rocket vehicle outside the earth's significant atmosphere between two fixed attitudes and in a fixed time.[3.4] An abrupt change in thrust level occurred at

a fixed time during flight. The time of thrust discontinuity determined the interface of the two subarcs which were optimized separately. Since the time of this discontinuity was fixed, t_o^2 could not be changed and the change in the state variables at the discontinuity were given by

$$\delta \underline{x}_o^2 = k\,(\underline{\psi}^2 - \underline{\psi}^1)\,, \quad k > 0 \tag{3.15}$$

(Here the sign preceding k is positive since the range is being maximized.) Thus, it is seen in this example that the 2nd-level gradient controller simply drives the trajectory at $t_f^1 = t_o^2$ to the point in space where the adjoint variables, $\underline{\psi}$, are continuous; a requirement of the Weierstrass – Erdmann corner conditions.

Computationally the subarcs were each minimized at the first level using the 2nd variational method of Breakwell.[3.5] The 2nd level gradient controller used Equation (3.15) and converged well until it approached the maximum range where the solution showed the expected oscillations about that maximum.

3.5 Conclusions

The multi-level optimization technique appears to be a straightforward way to synthesize interacting trajectories. The advantage of being able to use different optimization techniques on the different subarcs (provided the techniques will calculate the adjoint variables) may offset the disadvantage of repeated subarc optimizations. In any event this method allows one to concentrate effort on optimizing only small portions of the trajectory at a time (1st level). Then the 2nd-level controller can drive the first-level solutions to the overall optimal trajectory which may be highly complex.

4. VARIABLE TIME OPTIMAL TRAJECTORIES FOR AEROSPACE VEHICLES

4.1 Introduction

In problems of optimal trajectories, the final time T may be fixed or it may be variable. For each fixed time, an optimal trajectory may be determined which minimizes the given performance criterion. The variable time problem involves finding that final time for which the minimized performance criterion is minimal over all other final times.

The criteria used to determine the optimum in such techniques as Pontryagin's maximum principle are merely necessary conditions rather than sufficient conditions. In general it is assumed that solutions obtained using such necessary conditions are indeed optimum because if one knows that a minimum (or maximum) of the performance criterion exists, and that only one solution exists which satisfies the necessary conditions, then this solution must be the desired optimum. This assumption is generally valid for the fixed time problem when one is merely looking for a solution to a set of differential equations which satisfies given boundary conditions for a fixed time. However, in the variable time problem, the additional necessary condition that the Hamiltonian $H = 0$ may be satisfied for many different final times, as will be demonstrated in the sections to follow.

Let us consider the performance criterion given by:

$$J(u, T, X) = \int_0^T f_0[x(t),\ u(t)]dt \tag{4.1}$$

where $\underline{X} = \underline{x}(T)$ is the terminal state of the system given by $\dot{\underline{x}} = \underline{f}(x, u)$, with $\underline{u}$ the r-dimensional control vector. Let us define:

$$S(T, X) = J(\bar{u}, T, X) \tag{4.2}$$

where $\bar{u}$ is the optimal control which minimizes J for the given T and X.

It can be shown[7,9] that the Hamiltonian

$$H = -\frac{\partial S(T,X)}{\partial T} \tag{4.3}$$

which is, of course, the Hamilton-Jacobi equation from the calculus of variations. However, its relation to Pontryagin's necessary condition for a variable time optimum has never been properly explored.

If we consider a graph of $S(T,X)$ vs. T for fixed $\underline{X}$, that is, if we consider a graph of the values of the minimum performance criterion for each fixed time T plotted as a function of T, the Hamiltonian H is the negative of the derivative with respect to T of this curve. Thus the condition $H = 0$ for a minimum of $S(T,X)$ is a necessary condition in the same sense that the condition that the derivative of a general function be zero is a necessary condition for the minimum of that function. However, the condition $H = 0$ will thus also hold for a local maximum of the function $S(T,X)$, and any computational algorithm employing only the $H = 0$ condition for variable time optimum can yield such a local maximum if it exists. In the sections to follow, an example will be given in which both a local minimum and a local maximum of $S(T,X)$ do indeed exist (in fact, for which several such local maxima or minima exist).

The application of the maximum principle to a fixed time problem results in a two-point boundary value problem. A standard technique for solving the two-point boundary value problem is quasilinearization,[4,5] and this technique will be used in this paper to solve the fixed time problem. In order to solve the variable time problem, additional techniques are needed to vary the first time T so that the Hamiltonian is driven to zero. Such techniques will be developed in the sections to follow.

4.2 Computational Algorithm Using $H = 0$ Necessary Condition To Solve The Variable Time Problem

All methods of solving the variable time problem developed to this time use only the $H = 0$ necessary condition and assume the resulting solution is the only one and thus the desired optimum. One standard method is to iterate on the final time T and $n-1$ of the n initial adjoint conditions (n total variables) until the n terminal state variable conditions

are satisfied, and at each iteration use the H = 0 condition to determine the n^{th} initial adjoint condition. In this manner, the H = 0 condition is always satisfied, and when the boundary conditions are finally reached, the resulting solution is a possible solution of the variable time problem. However, in this process, the optimum solution at only one final time T is obtained.

Since we want to consider the variable time problem more or less separately from the other aspects of the optimal control problem in order to judge the validity of using only the H = 0 condition, we are going to derive an algorithm which solves a series of fixed time problems, iterating on the final time T until the H = 0 condition is satisfied. This algorithm will solve the fixed time problem for an initial guess on T, compute an increment on T, δT, such that when the fixed time problem for new final time $T + \delta T$ is solved, a Hamiltonian closer to zero will result, and through repeated iterations finally arrive at the T for which H = 0.

If we take our first solution for which the terminal conditions are satisfied at time T, and let the time vary by an amount δT, this will produce an error in terminal conditions which is given to a first order of approximation by

$$\delta \underline{x}(T) = \dot{\underline{x}}(T)\, \delta T \tag{4.4}$$

If we want to resolve the two-point boundary problem for the new $T^{n+1} = T^n + \delta T^n$, we will have to vary our initial adjoint condition $\underline{\lambda} = \underline{\psi}(0)$ by some vector $\delta \underline{\lambda}$. The question is what δT will require a $\delta \underline{\lambda}$ to correct the terminal conditions which will also make the Hamiltonian equal to zero. Since the Hamiltonian is constant, we need only require that it be zero at the initial conditions.

A standard set of differential equations may be written which may be used to calculate terms of the form $\partial x_i(t)/\partial \lambda_j$, which give the effect of varying the initial adjoint condition λ_j on the state variable $x_i(t)$. These equations are given by[4.4]

$$\frac{d}{dt}\left[\frac{\partial x}{\partial \lambda}\right] = \left[\frac{\partial^2 H}{\partial x \partial \psi}\right]\left[\frac{\partial x}{\partial \lambda}\right] + \left[\frac{\partial^2 H}{\partial \psi^2}\right]\left[\frac{\partial \psi}{\partial \lambda}\right]$$

$$\frac{d}{dt}\left[\frac{\partial \psi}{\partial \lambda}\right] = -\left[\frac{\partial^2 H}{\partial x^2}\right]\left[\frac{\partial x}{\partial \lambda}\right] - \left[\frac{\partial^2 H}{\partial \psi \partial x}\right]\left[\frac{\partial \psi}{\partial \lambda}\right] \qquad (4.5)$$

where $H = \underline{f}^T \cdot \underline{\psi} - f_o$, $\dot{\underline{\psi}} = -\left\{\frac{\partial H}{\partial x}\right\}$, $\left\{\frac{\partial H}{\partial x}\right\}$ is a n-vector with elements $\partial H/\partial x_i$ in, $[\partial \psi/\partial \lambda]$ and $[\partial x/\partial \lambda]$ are nxn matrices with elements $\partial \psi_i/\partial \lambda_j$ and $\partial x_i/\partial \lambda_j$, respectively. The initial conditions are $\partial x_i/\partial \lambda_j = 0$ for all i, j; $\partial \psi_i/\partial \lambda_j = 0$ for $i \neq j$; and $\partial \psi_i/\partial \lambda_j = 1$ for $i = j$. It should be noted that, in order to simplify (4.5), $u = u(x,\psi)$ has been substituted into $H = H(\psi, x, u)$ to make $H = H(\psi, x)$, thus eliminating terms such as $\left[\frac{\partial^2 H}{\partial \psi \partial u}\right]\left[\frac{\partial u}{\partial \lambda}\right]$ in (4.5). If we integrate these differential equations to time T, and let [Y] represent the matrix of elements $Y_{ij} = \partial x_i/\partial \lambda_j \,|\, t = T$, then a variation in initial adjoint conditions $\delta\underline{\lambda}$ will produce, to a first order of approximation, a change in the value of the state variables at time t = T given by $\delta\underline{x}(T) = [Y]\,\delta\underline{\lambda}$.

However, since the variation in $\underline{x}(T)$ due to a variation in final time of δT is given by (4.4), we want to vary $\underline{\psi}(0)$ by $\delta\underline{\lambda}$ to counteract this variation, that is, choose $\delta\underline{\lambda}$ so that

$$\delta\underline{x}(T) = -\dot{\underline{x}}(T)\,\delta T = [Y]\,\delta\underline{\lambda} \qquad (4.6)$$

Thus

$$\delta\underline{\lambda} = -[Y]^{-1}\,\dot{\underline{x}}(T)\,\delta T \qquad (4.7)$$

The Hamiltonian at t = 0 is given by

$$H^n = \underline{f}^T(0) \cdot \underline{\psi}^n(0) - f_o(0) \qquad (4.8)$$

where $f_o(0) = f_o\left(\underline{x}(0), \underline{\psi}(0)\right)$, and $\underline{f}(0) = \underline{f}\left(\underline{x}(0), \underline{\psi}(0)\right)$ since $\underline{u}(0) = \underline{u}\left(\underline{x}(0), \underline{\psi}(0)\right)$.

We want the new $H^{n+1} = \underline{f}^T(0) \cdot \underline{\psi}^{n+1}(0) - f_o(0) = 0$.

If we let $\underline{\psi}^{n+1}(0) = \underline{\psi}^{n} + \delta\underline{\lambda}^{n}$, and set $H^{n+1} = 0$,

$$H^{n+1} = f\left(\underline{x}(0), \underline{\psi}^{n}(0) + \delta\underline{\lambda}^{n}\right) \cdot \left(\underline{\psi}^{n}(0) + \delta\underline{\lambda}^{n}\right)$$

$$- f_{o}\left(\underline{x}(0), \underline{\psi}^{n}(0) + \delta\underline{\lambda}^{n}\right) = 0 \qquad (4.9)$$

If we substitute Equation (4.7) in (4.9), we obtain a single equation in a single unknown δT, which may be solved for δT. That is, [Y], $\dot{\underline{x}}(T)$, $\underline{\psi}^{n}(0)$, and $\underline{x}(0)$ are all either known originally or were determined in solving the problem for the previous T.

If we then apply the $\delta\underline{\lambda}$ given by (4.7) for the δT determined as above and apply $\underline{\psi}^{n+1}(0) = \underline{\psi}^{n}(0) + \delta\underline{\lambda}^{n}$ to the system of equations, we will find that the Hamiltonian is exactly equal to zero. However, when we integrate the system of equations, we will find that the terminal values are slightly off. This is due to the fact that $\delta\underline{x}(T)$ is not given exactly by (4.4), and even if it were, the $\delta\underline{\lambda}$ given by (4.7) would not exactly compensate for the $\delta\underline{x}(T)$ in one step. Thus we must again apply the fixed time algorithm to exactly solve the two-point boundary problem for the new $T^{n+1} = T^{n} + \delta T$. This will result in a $\underline{\psi}^{n+1}(0)'$ which is not exactly equal to $\underline{\psi}^{n+1}(0) = \underline{\psi}^{n}(0) + \delta\underline{\lambda}$ and thus the Hamiltonian will not be exactly zero. However, the $\psi^{n+1}(0)'$ determined by the fixed time algorithm for the T^{n+1} will be close enough to $\psi^{n+1}(0)$ that the Hamiltonian will be closer to zero than it was for the T^{n}. Thus, this process can be repeated until the T is found which does result in H = 0, or at least as close to zero as we please.

4.3 Numerical Example

The problem to be solved is given by the differential equations:[6]

$$\dot{x}_1 = x_2$$

$$\dot{x}_2 = (1-x_1^2)x_2 - x_1 + u \qquad (4.10)$$

with boundary conditions

$$x_1(0) = 1.0 \qquad x_1(T) = -0.97$$
$$x_2(0) = 0.0 \qquad x_2(T) = -0.96 \tag{4.11}$$

The performance criterion to be minimized was

$$J(u) = \int_o^T \left(x_1^2 + x_2^2 + u^2 \right) dt \tag{4.12}$$

The Hamiltonian for the system is thus given by

$$H = \psi_1 x_2 + \psi_2 \left[\left(1 - x_1^2 \right) x_2 - x_1 + u \right] - x_1^2 - x_2^2 - u^2 \tag{4.13}$$

If we set $\partial H/\partial u = 0$ and solve for u, we obtain

$$u_{opt} = \psi_{2/2} \tag{4.14}$$

If we have bounds on u of the form

$$u_{min} \leq u \leq u_{max} \tag{4.15}$$

we can see from (4.13) that if u_{opt} given by (4.15) is outside the bounds on u, then the Hamiltonian is maximized for admissible u if u is at its constraint. That is, if $u_{max} \leq \psi_{2/2}$, $u_{opt} = u_{max}$, and if $\psi_{2/2} \leq u_{min}$, $u_{opt} = u_{min}$.

The adjoint equations are given by

$$\dot{\psi}_1 = -\frac{\partial H}{\partial x_1} = 2x_1 x_2 \psi_2 + 2x_1 + \psi_2$$
$$\dot{\psi}_2 = -\frac{\partial H}{\partial x_2} = -\psi_1 - \psi_2 (1 - x_1^2) + 2x_2 \tag{4.16}$$

We will first solve the problem for unbounded control and then proceed on to the bounded control problem.

We have now reduced the problem to the two-point boundary problem and we can employ quasilinearization to solve this problem. After solving the problem for a given fixed time T, we employ the technique of the preceding section to calculate a δT to give us a new estimate $T + \delta T$. The equations for applying this technique are derived below.

From Equation (4.7), for a two-dimensional problem

$$\begin{Bmatrix} \delta\lambda_1 \\ \delta\lambda_2 \end{Bmatrix} = - \begin{bmatrix} \partial x_1/\partial\lambda_1 & \partial x_1/\partial\lambda_2 \\ \partial x_2/\partial\lambda_1 & \partial x_2/\partial\lambda_2 \end{bmatrix}^{-1}_{t=T} \begin{Bmatrix} \dot{x}_1(T) \\ \dot{x}_2(T) \end{Bmatrix} \delta T$$

$$= \begin{Bmatrix} R_1 \\ R_2 \end{Bmatrix} \delta T \qquad (4.17)$$

We of course must integrate Equations (4.5) to obtain the $\partial x_i/\partial\lambda_j$. R_1 and R_2 are given by

$$R_1 = -\left[\partial x_2/\partial\lambda_2 \Big|_{t=T} \dot{x}_1(T) - \partial x_2/\partial\lambda_1 \Big|_{t=T} \dot{x}_2(T) \right] / \det \left| \partial x/\partial\lambda \right|$$

$$R_2 = -\left[\partial x_1/\partial\lambda_1 \Big|_{t=T} \dot{x}_2(T) - \partial x_1/\partial\lambda_2 \Big|_{t=T} \dot{x}_1(T) \right] / \det \left| \partial x/\partial\lambda \right| \qquad (4.18)$$

where:

$$\det\left|\partial x/\partial\lambda\right| = \left(\partial x_1/\partial\lambda_1 \cdot \partial x_2/\partial\lambda_2 - \partial x_2/\partial\lambda_1 \cdot \partial x_1/\partial\lambda_2\right)\Big|_{t=T} \quad (4.19)$$

The Hamiltonian for our example, from (4.13), with $u = \psi_2/2$ substituted for unbounded control, is given by

$$H(\psi, x) = \psi_1 x_2 + \psi_2(1 - x_1^2)\, x_2 - \psi_2 x_1 + \psi_2^2/2 - x_1^2 - x_2^2 - \psi_2^2/4 \quad (4.20)$$

Since $H(\psi, x)$ is constant for all t, we need only insure that our $H(\psi + \delta\psi,\ x) = 0$ at $t=0$. Substituting the initial conditions $x_1(0) = 1.0$, $x_2(0) = 0$, and $\psi_2(0) = \lambda_2$

$$H(\lambda) = \lambda_2^2/4 - \lambda_2 - 1 \quad (4.21)$$

Therefore

$$\begin{aligned} H(\lambda + \delta\lambda) &= (\lambda_2 + \delta\lambda_2)^2/4 - (\lambda_2 + \delta\lambda_2) - 1 \\ &= \lambda_2^2/4 - \lambda_2 - 1 + \delta\lambda_2^2/4 + \frac{(\lambda_2 - 2)}{2}\,\delta\lambda_2 \\ &= H(\lambda) + \delta\lambda_2^2/4 + \frac{(\lambda_2 - 2)}{2}\,\delta\lambda_2 \end{aligned} \quad (4.22)$$

where $H(\lambda)$ and λ_2 are also available in computer storage after solution of the fixed time problem. Thus, substituting $\delta\lambda_2 = R_2\,\delta T$ from (32); and setting $H(\lambda + \delta\lambda) = H(\delta T) = 0$

$$\frac{R_2^2}{4}\,\delta T^2 + \frac{(\lambda_2 - 2)}{2} R_2 \delta T + H(\lambda) = 0 \quad (4.23)$$

This equation may be solved for δT, and thus $T + \delta T$ is our new estimate for the optimal final time. If (4.22) has no real solution, we use the δT that minimizes (4.22); if it has two real solutions, we choose the smaller, for which this approximations made hold more accurately.

For bounded control, if the control for the fixed time case is at its bound at $t = 0$, $H(\lambda + \delta\lambda)$ of (4.21) becomes

$$H(\lambda + \delta\lambda) = H(\lambda) - \delta\lambda_2(1-u)$$

$$= H(\lambda) - R_2(1-u)\delta T \tag{4.24}$$

and we use

$$\delta T = \frac{H(\lambda)}{R_2(1-u)} \tag{4.25}$$

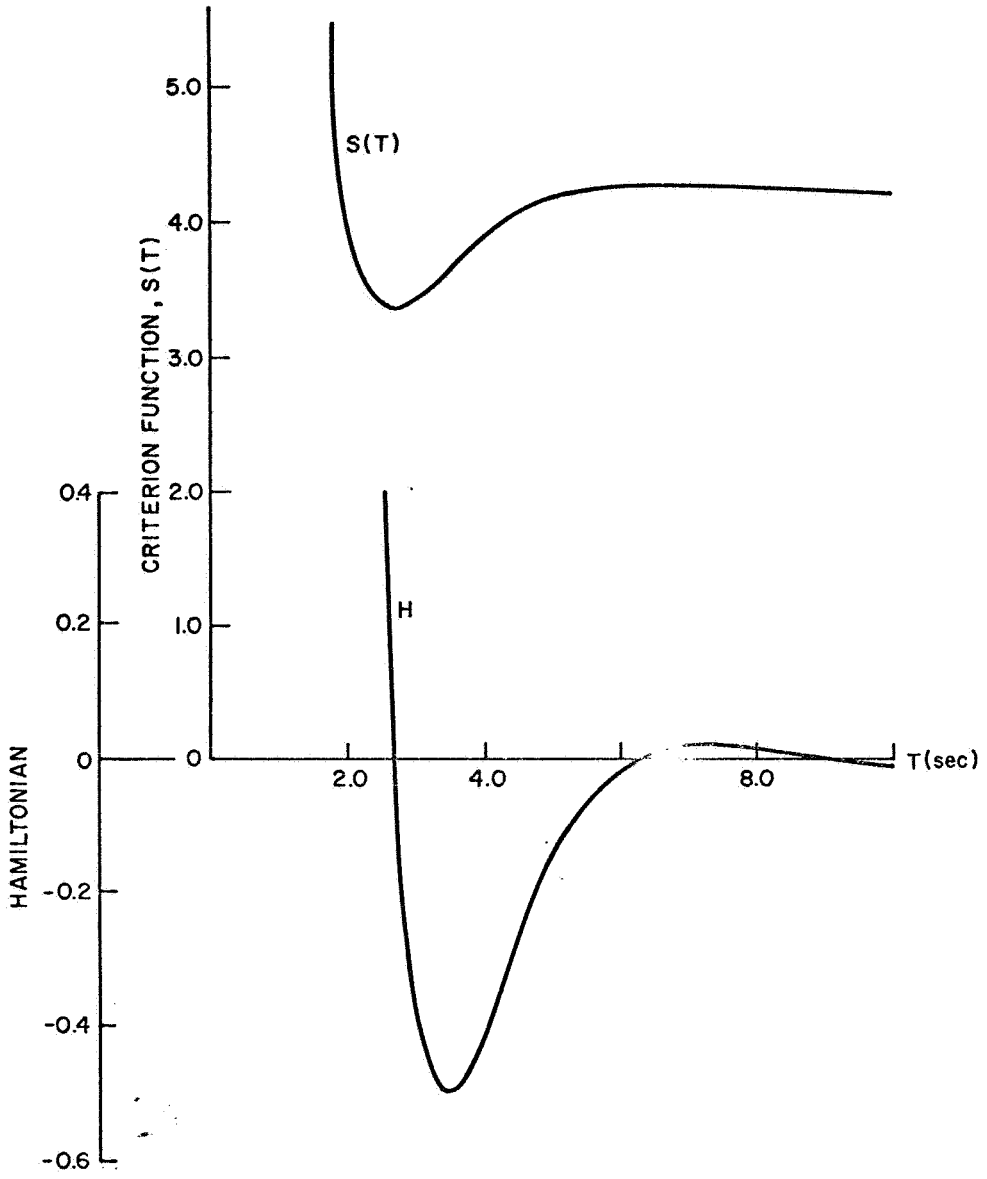
4.4 Discussion of Numerical Results

Figure 4.1 and Table 4.1 give the values of the Hamiltonian and the minimized performance criterion, S(T), for various values of fixed final time for unbounded control. As can be seen from the figure and table, the Hamiltonian goes through zero at $T \cong 2.6$ seconds, at which time S(T) goes through a minimum. This point is the desired variable time optimum. However, the Hamiltonian also goes through zero at $T \cong 6.1$ seconds, at which time S(T) attains a local maximum. Again, at $T = 9.3$ seconds, $H = 0$. According to the table, S(T) is not a local minimum at this point, but this is due to accumulated integration errors for large T's and the fact that S(T) is only varying in the fourth decimal place. S(T) should be a local minimum here (although a greater minimum than that attained at $T = 2.6$ seconds). For higher T's, H would probably continue to oscillate with decreasingly small amplitudes about $H = 0$, and S(T) would continue to achieve relative maxima and minima, but with increasingly small difference between them.

Any technique for solving the variable time problem which uses only the $H = 0$ criterion could arrive at any of the various $H = 0$ points, some being local maxima. For example, using the technique outlined in

TABLE 4.1

T (sec)	H	S(T)
2.0	2.8466	4.0320
2.2	1.3928	3.6217
2.4	0.5522	3.4347
2.6	0.0569	3.3780
2.8	-0.2338	3.3981
3.0	-0.3974	3.4627
3.2	-0.4792	3.5512
3.4	-0.5067	3.6502
3.6	-0.4981	3.7510
3.8	-0.4659	3.8475
4.0	-0.4194	3.9360
4.2	-0.3654	4.0144
4.4	-0.3088	4.0816
4.6	-0.2533	4.1376
4.8	-0.2013	4.1828
5.0	-0.1544	4.2181
5.2	-0.1134	4.2447
5.4	-0.0786	4.2636
5.6	-0.0498	4.2762
5.8	-0.0269	4.2837
6.0	-0.0090	4.2871
6.2	0.0041	4.2874
6.4	0.0135	4.2855
6.6	0.0195	4.2821
6.8	0.0229	4.2777
7.0	0.0243	4.2729
7.2	0.0240	4.2679
7.4	0.0227	4.2631
7.6	0.0207	4.2587
7.8	0.0184	4.2547
8.0	0.0156	4.2512
8.2	0.0131	4.2482
8.4	0.0104	4.2458
8.6	0.0080	4.2439
8.8	0.0059	4.2423
9.0	0.0025	4.2358
9.2	0.0010	4.2350
9.4	-0.0004	4.2345
9.6	-0.0013	4.2343
9.8	-0.0019	4.2343
10.0	-0.0033	4.2339



MINIMIZED CRITERIA FUNCTIONS
AND HAMILTONIAN VS. TIME

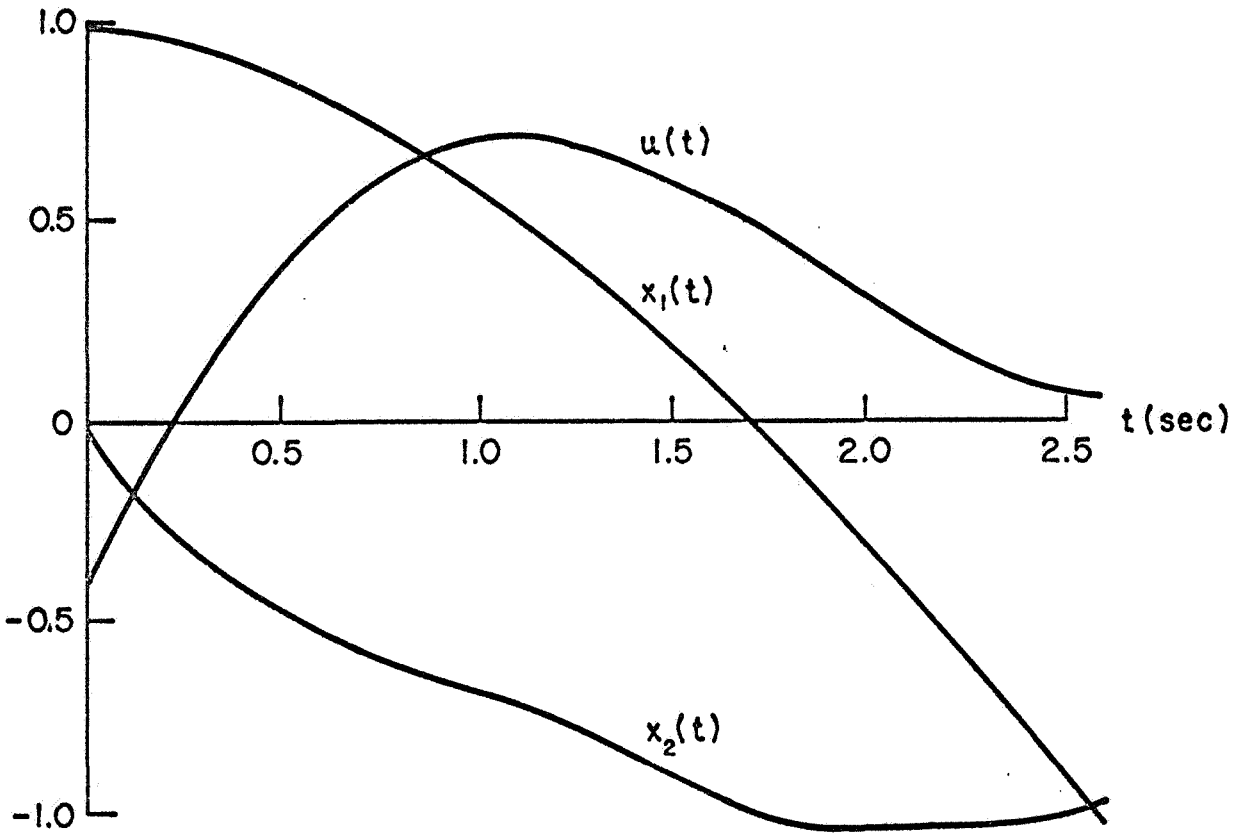
FIGURE 4. 1

preceding sections, if the initial guess on T was less than 3.4 seconds (the point where H achieves its first extremum), the variable time solution obtained will be the correct one of T = 2.6 seconds. However, if the initial guess on T is greater than 3.4 seconds (but less than seven seconds) the solution obtained will be the local maximum at T $\cong$ 6.1 seconds. This demonstrates the contention made earlier in this paper that all those methods given in the literature for solving the variable time problem which use only the H = 0 condition are inadequate for the general problem.

Using quasilinearization, and starting at T = 3.2 seconds, the variable time algorithm arrived at the solution T = 2.6 seconds after iterating through T = 2.0 and 2.4 seconds (or three total iterations on T). The variable time optimal trajectories for T = 2.6 seconds are shown in Figure 4.2. The first iteration overshot the desired T because the initial guess of T = 3.2 was near the extremum of H where H was varying slowly with respect to T, and thus the method computed a δT to drive H to zero which was too large. Thus, in order to make the technique work in regions where $\partial H/\partial T \cong 0$, an upper bound should be placed on the allowed δT(say, δT max = T/3).

4.5 Development of Sufficient Conditions for a Local Minimum

In the preceding section we showed that the H = 0 condition for a variable time optimum can sometimes lead to an incorrect solution. That is, computational algorithms based only upon the H = 0 condition can result in a solution which is a local maximum of S(T) with respect to T as well as a local minimum, since $H = -\frac{\partial S}{\partial T}$. In the standard minimization problem, a sufficient condition for a local minimum is that the second derivative of the function to be minimized be greater than zero. Since $H = -\frac{\partial S}{\partial T}$, for our problem this sufficient condition is given by:



VARIABLE TIME OPTIMAL TRAJECTORIES
UNBOUNDED CONTROL

FIGURE 4.2

$$\partial^2 S/\partial T^2 = -\frac{\partial H}{\partial T} > 0$$

or

$$\frac{\partial H}{\partial T} < 0 \tag{4.26}$$

Since the Hamiltonian is constant with time for the solution for a given final time T, $\partial H/\partial T$ will be the same regardless of at what time it is evaluated. For convenience we will evaluate it at time t = 0.

Since $H = H(\psi, x, u)$

$$\frac{\partial H}{\partial T} = \left\{\frac{\partial H}{\partial \psi}\right\}^T \left\{\frac{\partial \psi}{\partial T}\right\} + \left\{\frac{\partial H}{\partial x}\right\}^T \left\{\frac{\partial x}{\partial T}\right\} + \left\{\frac{\partial H}{\partial u}\right\}^T \left\{\frac{\partial u}{\partial T}\right\} \tag{4.27}$$

If we evaluate $\partial H/\partial T$ at t = 0, $\partial x/\partial T = 0$ since the initial condition is constant, and $\left\{\frac{\partial u}{\partial T}\right\} = \left\{\frac{\partial u(x,\psi)}{\partial T}\right\} = \left[\frac{\partial u}{\partial \psi}\right]\left\{\frac{\partial \psi}{\partial T}\right\}$ (where the term $\left[\frac{\partial u}{\partial x}\right]\left\{\frac{\partial x}{\partial T}\right\} = 0$ since $\left\{\frac{\partial x}{\partial T}\right\} = 0$).

Thus

$$\frac{\partial H}{\partial T} = \left\{\frac{\partial H}{\partial \lambda}\right\}^T \left\{\frac{\partial \lambda}{\partial T}\right\} + \left\{\frac{\partial H}{\partial u}\right\}^T \left[\frac{\partial u}{\partial \lambda}\right] \left\{\frac{\partial \lambda}{\partial T}\right\} \tag{4.28}$$

where $\psi = \lambda$ has been substituted since we are evaluating at t = 0.

From (4.7)

$$\delta\underline{\lambda} = -[Y]^{-1}\,\dot{\underline{x}}(T)\,\delta T$$

or

$$\frac{\delta\underline{\lambda}}{\delta T} = -[Y]^{-1}\,\dot{\underline{x}}(T) \tag{4.29}$$

But

$$\lim_{\delta T \to 0} \frac{\delta\underline{\lambda}}{\delta T} = \left\{\frac{\partial \lambda}{\partial T}\right\} \tag{4.30}$$

Therefore

$$\left\{\frac{\partial \lambda}{\partial T}\right\} = -[Y]^{-1}\,\dot{\underline{x}}(T) \tag{4.31}$$

Substituting (4.31) into (4.28)

$$\frac{\partial H}{\partial T} = -\left(\left\{\frac{\partial H}{\partial \lambda}\right\}^T + \left\{\frac{\partial H}{\partial u}\right\}^T \left[\frac{\partial u}{\partial \lambda}\right]\right) \cdot [Y]^{-1} \dot{\underline{x}}(T) \qquad (4.32)$$

In (4.32) all the variables on the right side of the equation may easily be evaluated for an optimal solution for variable final time obtained using the method of the previous chapter which uses only the $H = 0$ necessary condition. If $\frac{\partial H}{\partial T} < 0$, we have a local minimum; if $\frac{\partial H}{\partial T} > 0$ we have a local maximum.

We can also derive a computational algorithm for insuring that this sufficient condition is satisfied. If we follow the procedures developed earlier in this chapter, but include second order terms δT^2 is all approximation, we can obtain an expression for $\partial^2 H/\partial T^2$. Details of this development are given in Reference 4.7. Knowing $\partial^2 H/\partial T^2$, we know which way to vary T to make $\partial H/\partial T < 0$ if $\partial H/\partial T > 0$ for our first estimate on T. That is, we can take

$$\delta T = -K_T \frac{\partial H/\partial T}{\partial^2 H/\partial T^2} \qquad (4.33)$$

where the value of K_T depends on how far into the $\partial H/\partial T < 0$ region it is desired to drive T to insure convergence to a point where $\partial H/\partial T < 0$ (K_T should be at least two.) Of course, this method will only insure we have a local minimum and not a local maximum. If several local minima exist, we must search for the absolute minimum.

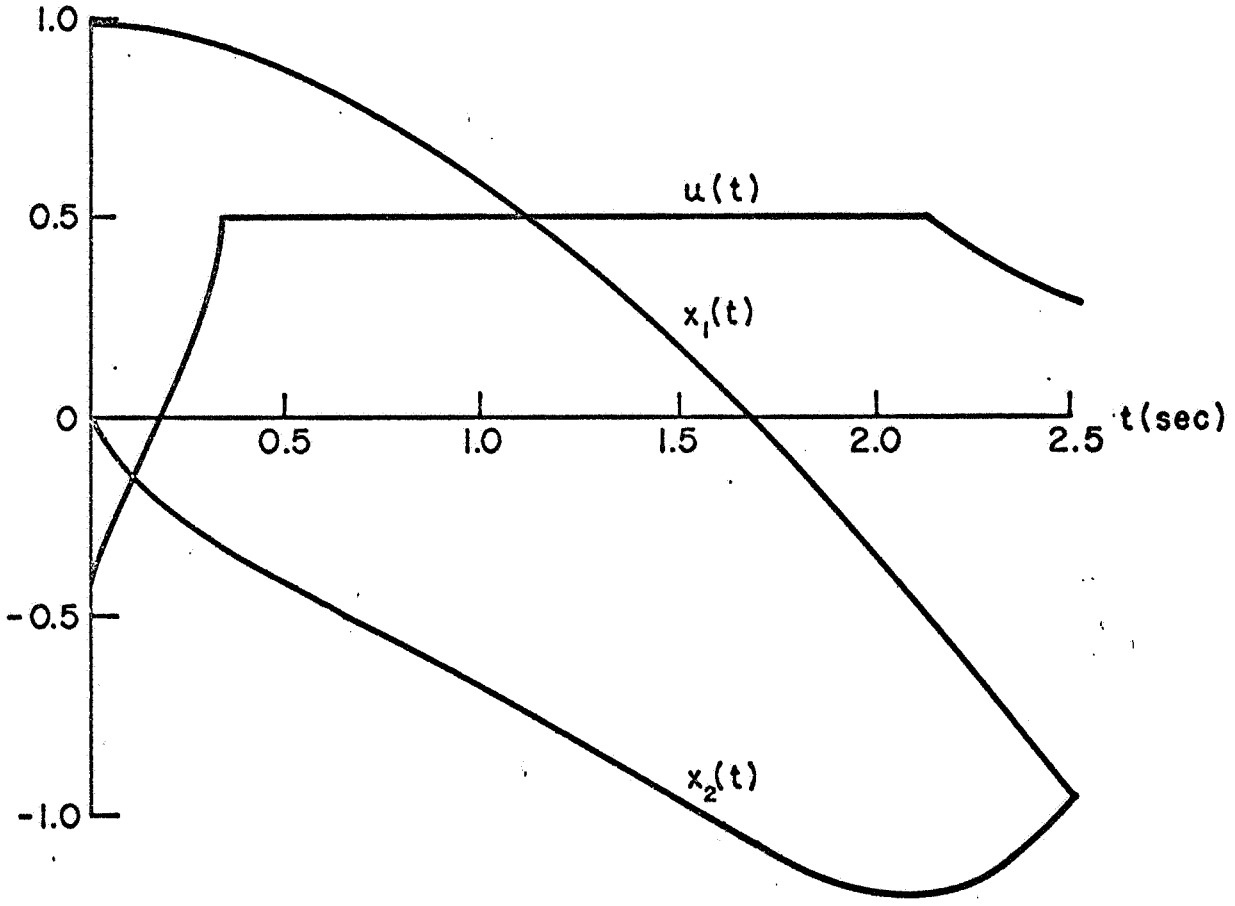
When this method was applied to the numerical example with an initial estimate of $T = 4.6$ seconds, the method converged to $T = 2.6$ seconds in steps of $T = 4.6$, 3.05, 2.2, 2.45, and 2.6 seconds, with $K_T = 2.0$ used for this run. The first iteration, as expected, drove T into the region where $\partial H/\partial T < 0$, and from there the method converged to the local minimum at $T = 2.6$ seconds. In the method using only the $H = 0$ criterion, any initial estimate of T greater than 3.4 seconds would converge to the local maximum of S(T) at $T = 6.1$ seconds.

When an initial estimate of $T = 5.2$ seconds was used (after $\partial^2 H/\partial T^2$ changes sign) the method converged to $T = 9.3$ seconds in steps

of 5.2, 7.8, 9.0, and 9.3 seconds. When only the H = 0 condition was used, the solution at T = 6.1 seconds was arrived at starting from T = 5.2 seconds. The solution at T = 9.3 seconds is a local minimum of S(T) (although Table 4.2 does not show this, due to integration errors), but not the least local minimum, which is at T = 2.6 seconds.

For the bounded control case, it was found that the most efficient way of solving the variable time problem was first to solve it for unbounded control and use the optimal final time T for unbounded control for the first estimate on T for the variable time optimal bounded problem.

If the bounded problem is solved in this manner, with bounds of ± 0.5, we obtain the variable time optimal solution at T = 2.525 seconds shown in Figure 4.4. As can be seen, even though the control is well bounded, the variable time optimal T is not too far from the value of T = 2.6 seconds for the unbounded case.



VARIABLE TIME OPTIMAL TRAJECTORIES
BOUNDED CONTROL

FIGURE 4.3

5. THE VARIABLE END POINT PROBLEM IN AEROSPACE VEHICLE TRAJECTORIES

5.1 Statement of the Problem

In the last section we considered the problem for which the terminal condition $\underline{x}(T)$ was some fixed constant vector $\{x_1(t), x_2(T), \ldots x_n(T)\}$ in the n-dimensional Euclidean space E_n. We will now consider problems in which the terminal vector $\underline{x}(T)$ is merely constrained to lie in some (n-k)-dimensional manifold M of E_n. Such a manifold is defined as the set of all points $\underline{x} = \{x_1, x_2 \ldots, x_n\}$ which satisfy simultaneously the k equations

$$\begin{aligned} h_1(x_1, x_2, \ldots, x_n) &= 0 \\ h_2(x_1, x_2, \ldots, x_n) &= 0 \\ h_k(x_1, x_2, \ldots, x_n) &= 0 \end{aligned} \tag{5.1}$$

and where the vectors

$$\text{grad } h_1(x), \text{ grad } h_2(x), \ldots \text{grad } h_k(x) \tag{5.2}$$

are linearly independent, where

$$\text{grad } h(x) = \left\{ \frac{\partial h}{\partial x_1}, \frac{\partial h}{\partial x_2}, \ldots \frac{\partial h}{\partial x_n} \right\} \tag{5.3}$$

Pontryagin[5.8] gives a necessary condition for the optimal control problem with fixed initial conditions and terminal condition constrained by $\underline{x}(T) \in M$. Let V be the tangent plane to the manifold M at a given point $\underline{x}$. This tangent plane has dimension (n-k). The desired additional necessary condition for this problem, inown as the transversality condition, is that for x(t) to be optimal, the terminal adjoint vector $\underline{\psi}(T)$ must be orthogonal to very vector $\underline{\theta} \in V$, i. e.,

$$\underline{\psi}^T(T) \cdot \underline{\theta} = 0 \tag{5.4}$$

for all $\underline{\theta} \in V$.

Since V is (n-k)-dimensional, (5.4) yields (n-k) independent relationships, which, along with the k equations (5.1) give sufficient relationships to solve the problem.

Let us consider S(T, X) as defined in the prefious chapter. The variable end point problem is to determine the minimum of S(T, X) over all $\underline{X}$ satisfying (5.1).

We will now derive the transversality condition by showing the equivalence between this condition and the condition that the first variation of S(T, X) with respect to $\underline{X}$ is zero, i.e.,

$$\delta S = \frac{\partial S}{\partial X}^{T} \cdot \delta \underline{X} = 0 \tag{5.5}$$

which is a necessary condition for S to be a minimum with respect to $\underline{X}$. $\frac{\partial S}{\partial X}$ is an n-dimensional vector with elements $\partial S/\partial X_i$ and $\delta \underline{X}$ is the n-dimensional variation in $\underline{X}$ such that $\underline{X} + \delta \underline{X}$ is also consistent with (5.1).

A standard result from optimal control theory is.[5.7, 5.9]

$$\frac{\partial S}{\partial X} = \underline{\psi}(T) \tag{5.6}$$

Thus, (5.5) becomes

$$\delta S = \underline{\psi}^{T}(T) \cdot \delta \underline{X} = 0 \tag{5.7}$$

However, we have stated that δX must be consistent with (5.1) i.e., to a first order approximation

$$\begin{aligned} \left\{\frac{\partial h_1}{\partial X}\right\}^{T} \cdot \delta \underline{X} &= 0 \\ \left\{\frac{\partial h_2}{\partial X}\right\}^{T} \cdot \delta \underline{X} &= 0 \\ &\cdots\cdots\cdots \\ \left\{\frac{\partial h_k}{\partial X}\right\}^{T} \cdot \delta \underline{X} &= 0 \end{aligned} \tag{5.8}$$

where $$\left\{\frac{\partial h_i}{\partial X}\right\} = \left\{\frac{\partial h_i}{\partial X_1}, \frac{\partial h_i}{\partial X_2}, \cdots \frac{\partial h_i}{\partial X_n}\right\}.$$

But the vectors $\{\partial h_i/\partial x\}$ are directed along the normals to the respective curves $h_i(x) = 0$, and thus the $\delta\underline{X}$ satisfying (5.8) are tangent to the respective curves $h_i(x) = 0$. Thus the $\delta\underline{X}$ satisfying all of (5.8) are tangent to all the curves $h_i(x) = 0$, and thus to the manifold M. Thus the $\delta\underline{X}$ in (5.7) are equivalent to the $\underline{\theta}$ in (5.4) (when $\delta\underline{X}$ goes to zero so that (5.8) become exact), demonstrating the equivalence between the transversality condition (5.4) and the condition $\delta S = 0$ of (5.5). Thus, the transversality condition for variable end point optimum is a necessary condition which holds for a variable end point maximum of S(T, X) as well as a minimum of S(T, X). Thus, if we use only this necessary condition, we can obtain a variable end point maximum as well as a minimum. Before commenting further on this, let us develop a computational algorithm using the transversality condition.

5.2 Computational Algorithm

We want to start with an initial estimate of $\underline{X}$ satisfying (5.1), solve the fixed end point problem as discussed earlier and then obtain a new estimate $\underline{X} + \delta\underline{X}$ so as to drive the condition (5.4) to zero.

We know that (5.4) gives us $r = n - k$ independent equations in the 2nd variables $x_1, x_2, \ldots x_n, \psi_1(T), \psi_2(T), \ldots \psi_n(T)$. Let the r equations be given by

$$\begin{aligned} M_1\big(X,\ \psi(T)\big) &= 0 \\ M_2\big(X,\ \psi(T)\big) &= 0 \\ \cdots\cdots\cdots\cdots & \\ M_r\big(X,\ \psi(T)\big) &= 0 \end{aligned} \tag{5.9}$$

Let us assume that our j^{th} estimate yields

$$M_i^j\left(X^j,\ \psi^j(T)\right) = \ell_i^j \tag{5.10}$$

We want to compute $\delta\underline{X}^{j+1}$ so that

$$M_i^{j+1}\left(X^j + \delta X^j, \psi^j(T) + \delta\psi^j(T)\right) = 0 \tag{5.11}$$

at least to a first order of approximation. Since $\underline{X}^j$ and $\underline{\psi}^j(T)$ are known from the previous trajectory, we have 2n unknowns

$\delta X_1, \delta X_2, \ldots, \delta X_n, \delta \psi_1(T), \delta \psi_2(T), \ldots \delta \psi_n(T)$.

And letting again, using (4.5),

$$Y_{ij} = \partial x_i / \partial \lambda_j \Big|_{t=T} \quad \text{and} \quad W_{ij} = \partial \psi_i / \partial \lambda_j \Big|_{t=T},$$

we can write that, to a first order of approximation, the variations in $\underline{X}$ and $\underline{\psi}(T)$ due to variations in $\underline{\lambda}$ are given by

$$\begin{aligned} \delta \underline{X} &= [Y]\, \delta \lambda \\ \delta \underline{\psi}(T) &= [W]\, \delta \underline{\lambda} \end{aligned} \tag{5.12}$$

If we combine Equations (5.12) we can obtain a relationship between $\delta \underline{\psi}(T)$ and $\delta \underline{X}$:

$$\delta \underline{\psi}(T) = [W]\, [Y]^{-1}\, \delta \underline{X} \tag{5.13}$$

Substituting (5.13) into (5.11) for $\delta \underline{\psi}(T)$, we reduce the number of unknowns in (5.11) from 2n to n. Also, the k Equations (5.8) can be used to eliminate k of the δX_i's, thus reducing the number of unknowns in (5.11) to $r = n - k$, which is equal to the number of equations. Thus (5.11) can be solved for (n-k) of the n δX_i's, giving (n-k) of the desired $X_i^{j+1} = X_i^j + \delta X_i$. The k Equations (5.1) can then be solved for the remaining k X_i^{j+1}'s, thus insuring that $\underline{X}^{j+1} \epsilon M$. Although, due to the first order approximations made, the $m_i^{j+1}\left(X, \psi(T)\right)$ will not exactly equal zero, they will be closer to zero than the previous $m_i^j\left(X, \psi(T)\right)$, and thus eventually the method will converge so that the $m_i\left(X, \psi(T)\right)$ are as close to zero as we please. This technique will be illustrated in the next section by a numerical example.

5.3 Numerical Example

We will use the same numerical example of the previous chapter, with terminal conditions constrained by

$$X_1^2 + X_2^2 = 0.5 \tag{5.14}$$

and with fixed final time $T = 5.0$ seconds and unbounded control.

The fixed end point problem for an initial estimate of $\underline{X}$ may be solved using quasilinearization. A new estimate $\underline{X} + \delta\underline{X}$ may be obtained by the following.

For the circle given by (5.14), the slope of the line drawn from the origin to a given point (X_1, X_2) is equal to X_2/X_1, and the slope of the perpendicular to this line, which is also tangent to the circle, is equal to $-X_1/X_2$. Thus the vector tangent to the circle at the given point is given by $(1, -X_1/X_2)$, or alternately, by $(-X_2, X_1)$.

From the transversality condition, we know that for $u(t)$ to be optimal for the variable end point problem, $\underline{\psi} = (\psi_1, \psi_2)$ must be perpendicular to this tangent at the final time $t = T$, that is,

$$\Big(\psi_1(T), \psi_2(T)\Big) \cdot (-X_2, X_1) = -\psi_1(T) \cdot X_2 + \psi_2(T) \cdot X_1 = 0 \tag{5.15}$$

Thus, we have $n-k = 2 - 1 = 1$ additional equation, which along with (5.14), is sufficient to solve the problem.

If we start with some initial estimate of the terminal conditions (X_1, X_2) consistent with (5.14) and solve the fixed end point problem, (5.15) will be equal to some value probably not equal to zero. We want to perturb $\underline{X}^n$ by some $\delta\underline{X}$ to get some $\underline{X}^{n+1} = \underline{X}^n + \delta\underline{X}$ so that (5.15) is equal to zero to a first order of approximation. Thus, we want

$$-(\psi_1 + \delta\psi_2)(X_2 + \delta X_2) + (\psi_2 + \delta\psi_2)(X_1 + \delta X_1) = 0 \tag{5.16}$$

From (5.13) we know that

$$\begin{bmatrix} \delta\psi_1(T) \\ \delta\psi_2(T) \end{bmatrix} = \begin{bmatrix} \partial\psi_1/\partial\lambda_1 & \partial\psi_1/\partial\lambda_2 \\ \partial\psi_2/\partial\lambda_1 & \partial\psi_2/\partial\lambda_2 \end{bmatrix}_{t=T} \begin{bmatrix} \partial x_1/\partial\lambda_1 & \partial x_1/\partial\lambda_2 \\ \partial x_2/\partial\lambda_1 & \partial x_2/\partial\lambda_2 \end{bmatrix}^{-1}_{t=T} \begin{bmatrix} \delta X_1 \\ \delta X_2 \end{bmatrix}$$

$$= \begin{bmatrix} K_{11} & K_{12} \\ K_{21} & K_{22} \end{bmatrix} \begin{bmatrix} \delta X_1 \\ \delta X_2 \end{bmatrix} \tag{5.17}$$

where, if we multiply out the matrices

$$
\begin{aligned}
K_{11} &= \left(\frac{\partial x_2}{\partial \lambda_2} \cdot \frac{\partial \psi_1}{\partial \lambda_1} - \frac{\partial x_2}{\partial \lambda_1} \cdot \frac{\partial \psi_1}{\partial \lambda_2}\right) \Big/ \det \left|\frac{\partial x}{\partial \lambda}\right| \\
K_{21} &= \left(\frac{\partial x_2}{\partial \lambda_2} \cdot \frac{\partial \psi_2}{\partial \lambda_1} - \frac{\partial x_2}{\partial \lambda_1} \cdot \frac{\partial \psi_2}{\partial \lambda_1}\right) \Big/ \det \left|\frac{\partial x}{\partial \lambda}\right| \\
K_{12} &= \left(\frac{\partial x_1}{\partial \lambda_2} \cdot \frac{\partial \psi_1}{\partial \lambda_1} + \frac{\partial x_1}{\partial \lambda_1} \frac{\partial \psi_1}{\partial \lambda_2}\right) \Big/ \det \left|\frac{\partial x}{\partial \lambda}\right| \\
K_{22} &= \left(\frac{\partial x_1}{\partial \lambda_2} \cdot \frac{\partial \psi_2}{\partial \lambda_1} + \frac{\partial x_1}{\partial \lambda_1} \cdot \frac{\partial \psi_2}{\partial \lambda_2}\right) \Big/ \det \left|\frac{\partial x}{\partial \lambda}\right|
\end{aligned}
\tag{5.18}
$$

where

$$
\det \left|\frac{\partial x}{\partial \lambda}\right| = \frac{\partial x_1}{\partial \lambda_1} \cdot \frac{\partial x_2}{\partial \lambda_2} - \frac{\partial x_2}{\partial \lambda_1} \cdot \frac{\partial x_1}{\partial \lambda_2}
$$

Also, from (5.14) and (5.8), to a first order of approximation, we have

$$
X_1 \delta X_1 + X_2 \delta X_2 = 0 \tag{5.19}
$$

Solving (5.19) for δX_1, substituting into (5.16) and (5.17), and then substituting (5.17) into (5.16) for $\delta \psi_1$ and $\delta \psi_2$ we obtain

$$
\begin{aligned}
&- \left(\psi_1 + K_{11} \delta X_1 - K_{12} \frac{X_1}{X_2} \delta X_1\right) \left(X_2 - \frac{X_1}{X_2} \delta X_1\right) \\
&+ \left(\psi_2 + K_{21} \delta X_1 - \frac{X_1}{X_2} K_{22} \delta X_1\right) \left(X_1 + \delta X_1\right) = 0
\end{aligned}
\tag{5.20}
$$

Collecting terms involving δX_1 and constants, and neglecting terms involving δX_1^2 since they are inaccurate anyway, we have an equation of the form

$$
A \, \delta X_1 + B = 0 \tag{5.21}
$$

where

$$A = \psi_1 \frac{X_1}{X_2} - X_2\left(K_{11} - K_{12}\frac{X_1}{X_2}\right) + \psi_2 + X_1\left(K_{21} - \frac{X_1}{X_2}K_{22}\right) \tag{5.22}$$

$$B = -X_2\psi_1 + X_1\psi_2$$

Thus

$$\delta X_1 = -B/A \tag{5.23}$$

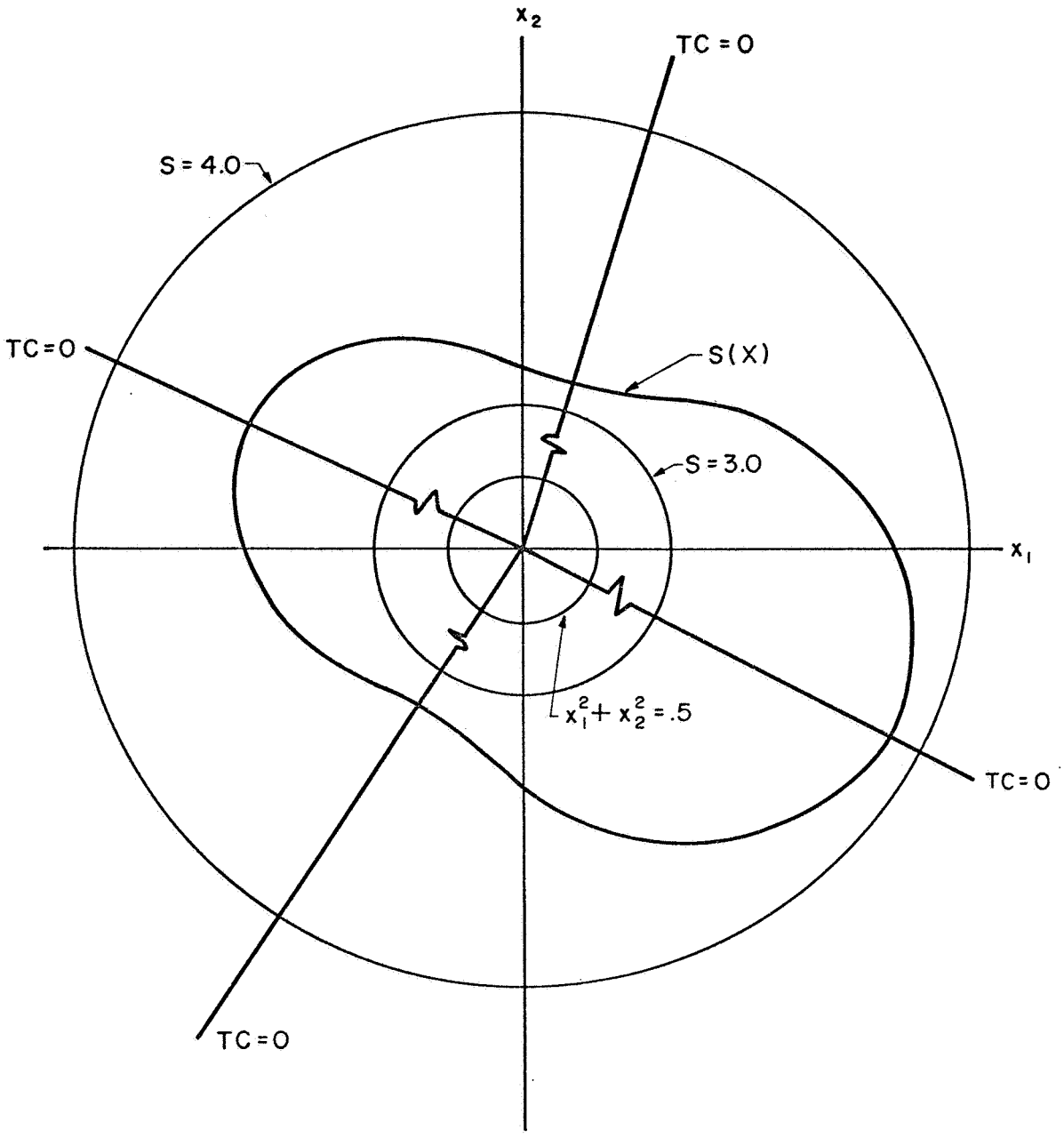
Also, we have

$$\delta X_2 = -\frac{X_1}{X_2}\,\delta X_1 \tag{5.24}$$

However, we cannot use both of these δX_i, since then (5.14) would not be satisfied exactly for $\underline{X}^{n+1} = \underline{X}^n + \delta\underline{X}$. Thus, if we take the smaller of the δX_i, and let $X_i^{n+1} = X_i^n + \delta X_i$ for that i, and then solve (5.14) for the other X_i^{n+1}, we will have our new estimate of $\underline{X}^{n+1}$ satisfying (5.14). We continue with this procedure until the transversality condition is as close to zero as we desire.

Numerical Results

Figure 5.1 gives a polar graph of the minimized criterion function $S = S(T, X)$ as a function of the end point $\underline{X}$ for a fixed final time $T = 5$ seconds. On this graph, S is plotted on lines radially outward from the origin with the S for a given $\underline{X}$ being plotted on the radial line through the given point on the $X_1^2 + X_2^2 = 0.5$ circle in the X_1-X_2 plane. The points on the circle where the transversality condition is zero are indicated by the radial lines marked TC = 0. As can be seen from this graph, there is a point where the transversality condition is zero in each of the four quadrants, with two points corresponding to local maxima and two to local minima. The local maxima occur in the second and fourth quadrants at $\underline{X}$ = (-0.64, 0.30) and (0.62, -0.35) with S = 3.51 and 3.85, respectively. The local minima occur in the first and third quadrants at $\underline{X}$ = (0.19, 0.68) and (-0.38, -0.60) with S = 3.08 and 3.14, respectively. Which of these extrema is obtained



MINIMIZED CRITERION FUNCTION
VS. END POINT

FIGURE 5.1

depends upon the first estimate of X used. Initial estimates were used at the intersection of the circle with the 45^{o} line in each of the four quadrants, i.e., at $\underline{X} = (\pm 0.50, \pm 0.50)$, and in each case the method converged to the extremum in the corresponding quadrant, e.g., an initial estimate of $(+0.5, +0.5)$ converged to the extremum at (0.19, 0.68).

5.4 Sufficient Conditions for a Local Minimum

If S(T, X) is at an extremum with respect to X, a sufficient condition for the extremum to be a minimum is that the second variation $\delta^2 S$ of S with respect to X be greater than zero. If $\delta S = \{\partial S/\partial X\}^T \cdot \delta \underline{X} = 0$, then the condition $\delta^2 S > 0$ is given by

$$\delta^2 S = \delta \underline{X}^T [\partial^2 S/\partial X^2] \, \delta \underline{X} > 0 \tag{5.25}$$

where $[\partial^2 S/\partial X^2]$ is an nxn matrix with elements $\partial^2 S/\partial X_i \partial X_j$. In (5.25), all the perturbations δX_i must be consistent with the constraints on $\underline{X}$, i.e., must satisfy Equations (5.8). Since (5.8) give us k relations between the δX_i, we may use (5.8) to eliminate k of the δX_i in (5.25) and thus we can reduce (5.25) to a relation consisting of (n-k)-dimensional vectors and a (n-k) X (n-k) matrix. Condition (5.25) is thus equivalent to the condition that the resulting (n-k) X (n-k) matrix is positive definite.

We now need a method of computing the $\partial S/\partial X_i \partial X_j$. From (5.6)

$$\left\{\frac{\partial S}{\partial X}\right\} = \underline{\psi}(T) \tag{5.26}$$

Therefore

$$\frac{\partial^2 S}{\partial X_i \partial X_j} = \frac{\partial \psi_i(T)}{\partial X_j} \tag{5.27}$$

We also know that

$$\delta \underline{\psi}(T) = \left[\frac{\partial \psi(T)}{\partial X}\right] \delta \underline{X} \tag{5.28}$$

Comparing (5.28) with (5.13), we see that

$$\left[\frac{\partial \psi(T)}{\partial X}\right] = [W]\,[Y]^{-1} \tag{5.29}$$

where the [W] and [Y] are defined as in (5.12). Thus this gives us all the information we need to compute $\delta^2 S$ and thus to apply the sufficient condition (5.25).

Let us illustrate this technique by applying it to our numerical example. From (5.29)

$$\frac{\partial \psi_i(T)}{\partial X_j} = K_{ij} \qquad i,j = 1,2 \tag{5.30}$$

where the K_{ij} are defined in (5.18). Thus for our two-dimensional problem, (5.25) becomes

$$\begin{Bmatrix} \delta X_1 \\ \delta X_2 \end{Bmatrix}^T \begin{bmatrix} K_{11} & K_{12} \\ K_{21} & K_{22} \end{bmatrix} \begin{Bmatrix} \delta X_1 \\ \delta X_2 \end{Bmatrix} > 0 \tag{5.31}$$

or

$$K_{11}\delta X_1^2 + \delta X_1\,\delta X_2(K_{12}+K_{21}) + K_{22}\,\delta X_2^2 > 0 \tag{5.32}$$

But, from the constraint Equation (5.19)

$$\delta X_1 = -\frac{X_2}{X_1}\,\delta X_2 \tag{5.33}$$

Thus (5.32) becomes

$$\left[\left(\frac{X_2}{X_1}\right)^2 K_{11} - \frac{X_2}{X_1}(K_{21}+K_{12}) + K_{22}\right]\delta X_2^2 > 0 \tag{5.34}$$

Thus the sufficient condition for a local minimum is that

$$\left(\frac{X_2}{X_1}\right)^2 K_{11} - \frac{X_2}{X_1}(K_{21}+K_{12}) + K_{22} > 0 \tag{5.35}$$

However, this condition will not insure that we have an absolute minimum, only that we have a local minimum. Thus this condition will not distinguish between the minima of the first and third quadrants of Figure 5.1.

5.5 Moving End Point Problem

In this section we will consider the class of problems in which end point is moving as a function of time. An example of this problem is that of reaching a point fixed on the surface of the rotating earth. This problem is neither a simple variable time problem, since the end point is varying, nor is it a variable time, variable end point problem in which the variations in final time and end point are independent and the $H = 0$ condition and the transversality condition can both be satisfied simultaneously. For the moving end point problem, we must derive a single necessary condition containing both the variations due to final time and end point. The first order variation in $S(T, X)$ for a variation in both T and $\underline{X}$ is given by

$$\delta S = - H \delta T + \left\{\frac{\partial S}{\partial X}\right\}^T \delta \underline{X} \tag{5.36}$$

since $H = - \partial S/\partial T$ (a standard result from the Hamilton-Jacobi equation).[5.7, 5.9] The end point $\underline{X}$ depends explicitly on the final time T. Let us represent this dependence by $\underline{X} = \underline{X}(T)$. (Note that $\underline{X}(T)$ is a vector which is a function of T rather than a vector evaluated at $t=T$.) Therefore, the variation $\delta \underline{X}$ is given by

$$\delta \underline{X} = \left\{\frac{dX}{dT}\right\} \delta T \tag{5.37}$$

where $\{dX/dT\}$ is an n-dimensional vector with components $dX_i(T)/dT$. Substituting (5.37) into (5.36)

$$\begin{aligned} \delta S &= - H \, \delta T + \left\{\frac{\partial S}{\partial X}\right\}^T \left\{\frac{dX}{dT}\right\} \delta T \\ &= \left(- H + \left\{\frac{\partial S}{\partial X}\right\}^T \left\{\frac{dX}{dT}\right\} \right) \delta T \end{aligned} \tag{5.38}$$

Dividing by δT and letting $\delta T \to 0$

$$\frac{dS}{dT} = - H + \left\{\frac{\partial S}{\partial X}\right\}^T \left\{\frac{dX}{dT}\right\} \tag{5.39}$$

The desired necessary for the moving end point problem is thus given by

$$- H + \left\{\frac{\partial S}{\partial X}\right\}^T \left\{\frac{dX}{dT}\right\} = 0 \tag{5.40}$$

This problem now becomes one of determing the value of final time T which drives (5.40) to zero.

In Section 4 a technique was developed for evaluating H at final time $T + \delta T$ given a value for H at final time T. This technique made use of a relationship giving the variation in $\underline{\lambda}$ caused by a variation in T. This equation is given by $\delta\underline{\lambda} = - [Y]^{-1} \dot{\underline{x}}(T)\, \delta T$. For the moving end point problem, this becomes

$$\delta \underline{\lambda} = - [Y]^{-1} \left(\dot{\underline{x}}(T) - \{dX/dT\}\right) \delta T \tag{5.41}$$

since the desired change in $\underline{X}$ is now $- \left(\dot{\underline{x}}(T) - \{dX/dT\}\right) \delta T$. Thus, since H is a constant function of t for a given T, H evaluated at t=0 is the same as H evaluated at t=T, and H evaluated at $T + \delta T$ may be obtained by taking $H(\underline{\lambda} + \delta\underline{\lambda})$ and substituting (5.41) for $\delta\underline{\lambda}$.

The relationship for the second term of (5.40) evaluated at $T + \delta T$ is given by

$$\left(\left\{\frac{\partial S}{\partial X}\right\}^T + \left\{\frac{dX}{dT}\right\}^T \left[\frac{\partial^2 S}{\partial X^2}\right] \delta T\right) \left(\left\{\frac{dX}{dT}\right\} + \left\{\frac{d^2X}{dT^2}\right\} \delta T\right) =$$

$$\left(\left\{\frac{\partial S}{\partial X}\right\}^T \left\{\frac{d^2X}{dT^2}\right\} + \left\{\frac{dX}{dT}\right\}^T \left[\frac{\partial^2 S}{\partial X^2}\right] \left\{\frac{dX}{dT}\right\}\right) \delta T \tag{5.42}$$

$$+ \left\{\frac{dX}{dT}\right\}^T \left[\frac{\partial^2 S}{\partial X^2}\right] \left\{\frac{d^2X}{dT^2}\right\} \delta T^2 \cdots$$

where $[\partial^2 S/\partial X^2]$ and $\{\partial S/\partial X\}$ are the same as given in (5.26) and (5.27), respectively. Combining this relationship with that for the estimate of H at $T + \delta T$, and setting it equal to zero as in (5.40), we get an equation which can be solved for the single unknown δT, which will drive (5.40) closer to zero. At each iteration we obtain the new value of the end point from $\underline{X} = \underline{X}(T + \delta T)$, our given function for the moving end point.

We will apply this technique to the same problem considered earlier in this chapter, with the initial condition $\underline{x}(0) = (1, 0)$, and the terminal condition moving around the circle $x_1^2 + x_2^2 = 0.5$ at a constant counter-clockwise angular velocity starting at the point (0.5, 0). Thus the terminal condition $\underline{X}(T)$ is given by

$$\begin{Bmatrix} X_1(T) \\ X_2(T) \end{Bmatrix} = \begin{Bmatrix} 0.5 \cos \omega\ t \\ 0.5 \sin \omega\ t \end{Bmatrix} \tag{5.43}$$

where ω is the constant angular velocity.

The various derivatives of $\underline{X}(T)$ with respect to T needed in (5.42) are given by

$$\begin{Bmatrix} \dfrac{dX_1}{dT} \\ \dfrac{dX_2}{dT} \end{Bmatrix} = \begin{Bmatrix} -0.5\ \omega\ X_2(T) \\ 0.5\ \omega\ X_1(T) \end{Bmatrix}$$

$$\begin{Bmatrix} \dfrac{d^2X_1}{dT^2} \\ \dfrac{d^2X_2}{dT^2} \end{Bmatrix} = \begin{Bmatrix} -0.5\ \omega^2 X_1(T) \\ -0.5\ \omega^2 X_2(T) \end{Bmatrix} \tag{5.44}$$

The $\{\partial S/\partial X\}$ and $[\partial^2 S/\partial X^2]$ are the same as those given in Section 5.4:

$$\left\{\frac{\partial S}{\partial X}\right\} = \underline{\psi}(T)$$
$$\left[\frac{\partial^2 S}{\partial X^2}\right] = [K] \qquad (5.45)$$

where the K_{ij} are defined in (5.18).

If we let the equation for (5.39) evaluated at $T + \delta T$ be given by $A\,\delta T^2 + B\,\delta T + C = 0$, the values of A, B, and C for our numerical example are given by

$$A = -\frac{\left(R_2'\right)^2}{\psi} + \begin{Bmatrix} -.5\,\omega\,X_2(T) \\ .5\,\omega\,X_1(T) \end{Bmatrix}^T \begin{bmatrix} K_{11} & K_{12} \\ K_{21} & K_{22} \end{bmatrix} \begin{Bmatrix} -.5\,\omega^2 X_1(T) \\ -.5\,\omega^2 X_2(T) \end{Bmatrix}$$

$$B = -\frac{\left(\lambda_2 - 2\right)}{2} R_2' + \begin{Bmatrix} \psi_1(T) \\ \psi_2(T) \end{Bmatrix}^T \begin{Bmatrix} -.5\,\omega\,X_2(T) \\ .5\,\omega\,X_1(T) \end{Bmatrix} \qquad (5.46)$$
$$+ \begin{Bmatrix} -.5\,\omega\,X_2(T) \\ .5\,\omega\,X_1(T) \end{Bmatrix} \begin{bmatrix} K_{11} & K_{12} \\ K_{21} & K_{22} \end{bmatrix} \begin{Bmatrix} -.5\,\omega\,X_2(T) \\ .5\,\omega\,X_1(T) \end{Bmatrix}$$

$$C = -H + \begin{Bmatrix} \psi_1(T) \\ \psi_2(T) \end{Bmatrix}^T \begin{Bmatrix} -.5\,\omega\,X_2(T) \\ .5\,\omega\,X_1(T) \end{Bmatrix}$$

where R_2' is the same as that given in (4.18) with $\dot{x}_1(T)$ and $\dot{x}_2(T)$ being replaced by $\dot{x}_1(T) + 0.5\,\omega\,\dot{x}_2(T)$ and $\dot{x}_2(T) - 0.5\,\omega\,\dot{x}_1(T)$, respectively. All the other terms in (5.46) are evaluated at final time T for the previous iteration and are available in computer storage. Thus the equation $A\,\delta T^2 + B\,\delta T + C = 0$ canbe solved for δT.

The value of ω used was such as to require 20 seconds for the end point to move completely around the circle, i.e., $\omega = 2\pi/20$. Starting from an initial estimate of T = three seconds, the method converged in three iterations to a local minimum at T = 4.75 seconds, $\underline{X} = (0.036, 0.499)$,

with S = 2.975. Starting from T = 6.5 seconds, the method converged in two iterations to a local maximum at T = 8.4 seconds, $\underline{X}$ = (-0.438, 0.241), with S=3.271. Starting from T=12 seconds, the method converged in 3 iterations to a local minimum at T = 13.5 seconds, $\underline{X}$ = (-0.216, -0.451), with S = 2.992. Note that in this case, there is very little difference between the two local minima. It is likely that S would continue to go through local maxima and minima as T continues to increase.

Sufficient conditions for a local minimum may be developed by differentiating (5.39) with respect to T

$$\frac{d^2S}{dT^2} = -\frac{dH}{dT} + \left\{\frac{\partial S}{\partial X}\right\}^T \left\{\frac{d^2X}{dT^2}\right\} + \left\{\frac{dX}{dT}\right\}^T \left[\frac{\partial^2 S}{\partial X^2}\right] \left\{\frac{dX}{dT}\right\} \tag{5.47}$$

A relationship for dH/dT may be obtained from (4.32) by replacing $\dot{\underline{x}}(T)$ by $\dot{\underline{x}}(T) - \{dX/dT\}$

$$\frac{dH}{dT} = -\left(\left\{\frac{\partial H}{\partial \lambda}\right\}^T + \left\{\frac{\partial H}{\partial u}\right\}^T \left[\frac{\partial u}{\partial \lambda}\right]\right) \cdot [Y]^{-1} \left(\dot{\underline{x}}(T) - \left\{\frac{dX}{dT}\right\}\right) \tag{5.48}$$

All other expressions in (5.46) have already been evaluated. The sufficient condition for a local minimum then is that $d^2S/dT^2 > 0$.

REFERENCES FOR SECTION 2

2.1 McGill, R. and P. Kenneth, "A Convergence Theorem on the Iterative Solution of Nonlinear Two Point Boundary Value Systems," XIV International Astronautical Congress, Proceedings, International Astronautic Federation, September 1963, Paris, France. Gauthier-Villars, Paris, 1965, Vol. 4, pp. 173-188.

2.2 McGill, R. and P. Kenneth, "Solution of Variational Problems by Means of a Generalized Newton-Raphson Operator," Grumman Aircraft, Engineering Report, GRD-100A, April 1964.

2.3 Bellman, Robert and Robert Kalaba, Quasilinearization and Nonlinear Boundary Value Problems, American Elsevier Publishing Company, New York, 1965.

2.4 McGill, R., "Optimal Control, Inequality State Constraints and the Generalized Newton-Raphson Algorithm," SIAM Journal on Control, Series A, Vol. 3, No. 2, pp. 291-298, December 1965.

2.5 Kopp, R. and H.G. Moyer, "Trajectory Optimization Technique," Advances in Control Systems, Vol. III, Ed., C.T. Leondes, Academic Press, New York, 1966, pp. 69-109.

2.6 Paine, G., "The Application of the Method Quasilinearization to the Computation of Optimal Control," Ph. D. Dissertation, Department of Engineering, University of California, Los Angeles, September 1966.

REFERENCES FOR SECTION 3

3.1 Macko, D. and J.D. Pearson, "A Multi-level Formulation Of Nonlinear Dynamic Optimization Problems in Papers on Multi-level Control Systems," Report No. SRC 70-A-65-25, Systems Research Center, Case Institute of Technology, Cleveland, Ohio, 1965.

3.2 Lasdon, L.S. and J.D. Schoeffler,"A Multi-level Technique For Optimization," Proc. 1965 Joint Autom. Control Conf., Rensselaer Polytechnic Institute, Troy, New York.

3.3 Brosilow, C.B., L.S. Lasdon and J.D. Pearson, "Feasible Optimization Methods For Interconnected Systems, Proc. 1965 Joint Autom. Control Conf., Rensselaer Polytechnic Institute, Troy, New York.

3.4 Bauman, E.J., Multi-level Optimization Techniques With Application to Trajectory Decomposition. Ph.D. Dissertation, Dept. Eng., University of Calif., Los Angeles, Calif., 1966.

3.5 Breakwell, J.V. and others, "Optimization and Control of Nonlinear Systems Using the Second Variation," Journal of Control, Ser. A, Vol. I, No. 2 (1963).

REFERENCES FOR SECTION 4

4.1 Breakwell, J., J. Speyer and A. Bryson, "Optimization and Control of Nonlinear Systems Using the Second Variation," SIAM Journal on Control, 1, No. 2, 193-223 (1963).

4.2 Bryson, A., W. Denham, F.J. Carroll and M. Kinya, "Determination of Lift or Drag Programs to Minimize Re-Entry Heating," Journal of the Aerospace Sciences, pp. 420-430, April 1962.

4.3 Gelfand, I.M., and S.V. Fomin, Calculus of Variations, Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1963.

4.4 Isaacs, D., An Algorithm for Sequential Optimization of Control Systems, Ph.D. Dissertation, Department of Engineering, University of California, Los Angeles, January 1965.

4.5 McGill, R., and P. Kenneth, "Solution of Variational Problems by Means of a Generalized Newton-Raphson Operator," Grumman Aircraft, Engineering Report, to appear in 1966.

4.6 Merriam, C.W., Optimization Theory and the Design of Feedback Control Systems, McGraw-Hill Book Company, New York 1964.

4.7 Niemann, Robert A., Variable Time Optimal Control, Ph.D. Dissertation, Department of Engineering, University of California, Los Angeles, December 1966.

4.8 Pontryagin, L.S., et al., The Mathematical Theory of Optimal Processes, Wiley and Sons, New York, 1962.

4.9 Snow, D.R., Reachable Regions and Optimal Controls, Ph.D. Dissertation, Department of Mathematics, Stanford University, October 1964.

REFERENCES FOR SECTION 5

5.1 Breakwell, J., J. Speyer and A. Bryson, "Optimization and Control of Nonlinear Systems Using the Second Variation," SIAM Journal on Control, 1, No. 2, 193-223 (1963).

5.2 Bryson, A., W. Denham, F.J. Carroll and M. Kinya, "Determination of Lift or Drag Programs to Minimize Re-Entry Heating," Journal of the Aerospace Sciences, pp. 420-430, April 1962.

5.3 Gelfand, I.M., and S.V. Fomin, Calculus of Variations, Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1963.

5.4 Isaacs, D., An Algorithm for Sequential Optimization of Control Systems, Ph.D. Dissertation, Department of Engineering, University of California, Los Angeles, January 1965.

5.5 McGill, R., and P. Kenneth, "Solution of Variational Problems by Means of a Generalized Newton-Raphson Operator," Grumman Aircraft, Engineering Report, to appear in 1966.

5.6 Merriam, C.W., Optimization Theory and the Design of Feedback Control Systems, McGraw-Hill Book Company, New York 1964.

5.7 Niemann, Robert A., Variable Time Optimal Control, Ph.D. Dissertation, Department of Engineering, University of California, Los Angeles, December 1966.

5.8 Pontryagin, L.S., et al., The Mathematical Theory of Optimal Processes, Wiley and Sons, New York, 1962.

5.9 Snow, D.R., Reachable Regions and Optimal Controls, Ph.D. Dissertation, Department of Mathematics, Stanford University, October 1964.