I. INTRODUCTION

When controlling the performance of a system, it is often desirable to choose the control that will minimize errors in the system and do it in the shortest possible time. A practical matter that must be considered in the optimization in relation to rapid action is the fact that control is of a bounded nature. In a great many important cases, the constraint on the magnitude of the control effort precludes the use of classical variational techniques to design the controller.

In 1956 Pontryagin hypothesized his "maximum principle" which has since been proven a necessary condition for the optimization of linear systems in relation to rapid action. In solving the minimum time problem for linear systems with bounded control, the principle leads to a "bang-bang" form of control law. This implies that the control effort is always being applied at its maximum value. There remains, however, the task of finding the optimum time to switch the control. Pontryagin's method leads to a rule for switching the controller which is a function of the initial conditions in the system adjoint to the one being controlled. Generally these initial conditions are difficult to find.

It is usually helpful to consider the control problem using state space techniques. The coordinates of the space for an nth order system here are a displacement error and its n-1 time derivatives. The space may be divided into two regions each of which is characterized by the control optimal for the trajectories in that region. Optimum switching between the two conditions of the bang-bang control occur on the hypersurface dividing the space. The switching criteria can then be stated as a function of the state space variables.

A problem of interest occurs when the system is of such a nature that when control is applied...
plied, a discontinuity appears in one or more of the system states. This may happen when the control is of a bang-bang form and the forward transmission path of the system contains zeros. It could also show up if the control is of such a form that it approximates an impulse to the system. When there are discontinuities in the state space due to switching it is generally no longer possible to write the switching criteria as a function of the state space variables.

One alternative might be to switch the control as a function of time. This may be done effectively when the number of switchings to reach the origin of the error state space is no more than \( n-1 \) in an \( n^{th} \) order system. Such a restriction limits one mainly to considering only those systems with real, distinct eigenvalues. Large disturbances in lightly damped (oscillatory) systems may require more than \( n-1 \) switchings to zero the error states. The most important consideration when controlling as a function of time is the means of implementing the switching logic. To accomplish time dependent control, it is virtually mandatory that a digital computer be inserted in the control loop.

Another approach to the problem is to find a system that reacts identically to the system with zeros except at the points of discontinuity. Control of this parallel system can be stated in terms of the state space variables. This logic can then be used to switch the original plant.

This paper will be an investigation into the latter method. The problem is as follows:

Given a second order oscillatory system with one zero, find the optimum control for zeroing the errors in the system in minimum time and for zeroing the errors with minimum fuel.

The method of Pontryagin is used to solve the problem. The brief description of the method presented here is based on the work of Rozonoer.\(^1\)

II. PONTRYAGIN'S MAXIMUM PRINCIPLE

Given the system state variables described by \( n \) first order differential equations

\[
x_i = f_i(x,u,t) \quad i=1, \ldots, n
\]

where \( x \) is a column vector in phase space and \( u \) is a column control vector consisting of \( r \) control elements.

The control \( u(t) \) must belong to a closed subset \( U \) of admissible controls and must be piecewise continuous. The trajectory \( x(t) \) in the phase space is uniquely determined by (1) when control \( u(t) \) and the initial conditions

\[
x(0) = x^0 = \begin{bmatrix} x^0_1 \\ \vdots \\ x^0_n \end{bmatrix}
\]

are given.

The control \( u(t) \) of a system may be considered optimum under a variety of criteria. A large class of optimization problems may be solved by presenting the criteria in such a way that the solution is attained by minimizing a linear function of the final value of the state space variables. A control must be selected from \( U \) that will transfer the system (1) from \( x^0 \) to some fixed closed set \( G \) of the phase space such that

\[
S = \sum_{i=1}^{n+1} c_i x_i(T)
\]

is a minimum. The constants \( c_i \) and the \( x_{n+1} \) coordinate are chosen such that minimizing (3) optimizes the system.

In a great many cases optimization of only one of the coordinates of the system is desired. For example, in order to optimize the magnitude of

\[
\int_0^T F(x(t),u(t)) \, dt
\]

for \( T \) and \( x(T) \) either fixed or free in a system (1) for \( u(t), U \), a new variable is introduced:

\[
x_{n+1} = \int_0^t F(x(t),u(t)) \, dt \quad \text{and another differential equation}
\]

\[
x_{n+1} = F(x(t),u(t))
\]

is added to (1). The problem of optimizing the integral leads to optimizing \( x_{n+1}(T) \) at \( t=T \).
Minimizing \( x_{n+1}(T) \) in the system (1) with \( x_{n+1}(t) \) adjoined is accomplished by putting the problem in functional form (3) and applying the maximum principle to gain the solution. That is

\[
S = \sum_{i=1}^{n+1} c_i x_i(T) = x_{n+1}(T)
\]

(6)
is the functional to be minimized. Here it may be seen that \( c_1 = c_2 = \ldots, c_n = 0 \) and \( c_{n+1} = 1 \).

A new dependent variable \( p(t) \) is now formed such that

\[
p_i(t) = - \sum_{i=1}^{n+1} p_i \frac{\partial f_i(x,u,t)}{\partial x_i}
\]

(7)
The function

\[
H = \sum_{i=1}^{n+1} p_i f_i(x,u,t)
\]

(8)
is introduced from which equations (1) and (7) may now be written

\[
x_i = \frac{\partial H}{\partial p_i}, \quad p_i = - \frac{\partial H}{\partial x_i}, \quad i = 1, \ldots, n+1
\]

(9)
The control \( u^*(t) \) is said to satisfy the maximum condition if \( H(x^*(t),p^*(t),u^*(t)) \) reaches an absolute maximum at each time \( t \) \((0 \leq t \leq T)\) where \( x^*(t) \) and \( p^*(t) \) are the values of the variables at time \( t \) with \( u^*(t) \in U \) controlling. For linear systems of the type discussed in this paper, the necessary and sufficient condition for minimizing

\[
S = \sum_{i=1}^{n+1} c_i x_i(T)
\]
optimally with admissible control is that the control satisfy the maximum condition.

To use the maximum principle, \( H \) is formed and maximized with respect to \( u(t) \). This produces a

\[
u^*(t) = \phi(x,p)
\]

(10)
which may be used with Equations (9) and the boundary conditions to find \( u^*(x) \). If the end point of \( x(t) \) is not fixed, it becomes necessary to obtain boundary conditions on \( p(t) \) in order to arrive at a solution. The conditions \( p(T) \) may be found using a function \( F(x) \equiv 0 \) which describes \( G \) and \( x^1(T) \in G \), the end point of an optimum trajectory. The form of \( p(T) \) will be stated without detailed explanation; however, it may be noticed that at time \( t = T \), \( p(T) \) is orthogonal to a hyperplane

\[
\sum_{i=1}^{n+1} a_i (x_i - x^1_i) = 0
\]

through the endpoint of the trajectory and directed toward that portion of \( G \) where

\[
\sum_{i=1}^{n+1} c_i x_i \leq \sum_{i=1}^{n+1} c_i x_i(T).
\]
The coefficients \( a_i \) may be expressed as a linear combination of the \( c_i \) and \( b_i(x^1(T)) \), the latter being coefficients of a hyperplane through \( x^1(T) \) bracketting \( G \). Thus

\[
p_i(T) = - \lambda c_i - \mu b_i(x^1(T))
\]

(11)
where \( \lambda \) and \( \mu \) are non-negative numbers one of which may be set equal to unity as it is only the ratio that is important.

Generally, three situations arise as to final boundary conditions.

(i) If \( x_i(T) \) are specified for \( i = 1,2,\ldots,m \) then these become the boundary conditions for (9).

(ii) If \( x_i^1(T) \) are internal points of \( G \) for \( i(1 \leq i \leq n+1) \) then \( b_i(x^1(T)) = 0 \) and \( p_i(T) = - c_i \).

(iii) If \( x_i^1(T) \) are boundary points of \( G \) for some \( i(1 \leq i \leq n+1) \) then \( F(x(T)) = 0 \) and the \( p_i(T) \) are as in (11).

When \( F \) is differentiable, the bracketting hyperplane through \( x^1 \) has coefficients

\[
b_i x^1(T) = \frac{\partial F}{\partial x_i} \bigg|_{x_i = x_i(T)}
\]

(12)
If finding the optimum control for minimum transit time another condition must be fulfilled since \( T \) is not fixed beforehand. This condition is that \( H(T) = 0 \).
III. DEVELOPMENT OF SYSTEM EQUATIONS

The equation of a second order system with zeros may be written

$$\ddot{c} + 2\zeta \omega c + \omega^2 c = a_1 u + a_2 u$$

where $c$ is the output variable of the system and $u$ is the output of a controller.

This paper is concerned with control of similar systems that are purely oscillatory in nature, i.e., $E_\eta = 0$. To facilitate ease of computation in the analysis, Equation (13) is scaled to

$$c + c' = a_1 u + u$$

which when written in terms of the Laplace transform of the output variable becomes

$$C(s) = \frac{(a_1 s + 1) U(s)}{s^2 + 1}$$

This system is represented in block diagram form in Fig. 1.

![Figure 1. Block Diagram of Control System.](image)

The response of the system to a step input is investigated more readily by means of the error variable

$$e = r - c$$

If the input $r$ is fed forward as in Fig. 2, the Laplace transform of the error, given

![Figure 2. Controlled System with Input Fed Forward.](image)

becomes

$$E(s) = \frac{(r_0 - c_0)s - (c_0 + a_1 r_0) - (a_1 s + 1) U(s)}{s^2 + 1}$$

Now the problem of zeroing the error states reduces to that of zeroing the error initial conditions in the system.

Finally with the introduction of state space variables

$$e_1 = -e$$
$$e_2 = e$$

the system equations can be written in vector matrix notation

$$e = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} e + \begin{bmatrix} 0 \\ a_1 u + u \end{bmatrix}$$

IV. THE MINIMUM TIME PROBLEM

The problem is stated as follows:

Given the system (20) and a control force of bounded magnitude $|u| \leq 1$, find the optimum control $u^*(t)$ to transfer the state variables from some initial point in the phase space to the origin of the phase space in minimum time $T$.

That is, given

$$e(0) = e^o$$
$$e(T) = 0$$

and the system (20), find $u^*(t)$ such that

$$T = \int_0^T \alpha \, dt$$

is a minimum where $\alpha$ is a positive constant.

Introduce

$$e_{n+1} = e_\eta = S = \int_0^T \alpha \, dt$$

The system equations then become

$$e = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} e + \begin{bmatrix} 0 \\ a_1 u + u \end{bmatrix}$$

Because of (21), the functional

$$S = \sum_1^3 c_\eta e(T) = c_\eta e_\eta(T)$$
and since we wish to minimize this, \( c_3 = 1 \) is chosen. \( e_3(T) \) is not limited, hence the boundary condition becomes

\[
p_3(T) = -c_3 = -1 \tag{26}
\]

By (8), the hamiltonian becomes

\[
H = p_1 e_2 - p_2 e_1 + p_2 (a_1 u + u) + p_3 \alpha \tag{27}
\]

Since

\[
p_3 = -\frac{\partial H}{\partial e_3} = 0
\]

it is evident that \( p_3 \) is a constant and therefore \( p_3 = p_3(T) = 1 \) and now \( H \) is

\[
H = p_1 e_2 - p_2 e_1 + p_2 (a_1 u + u) - \alpha \tag{28}
\]

which is maximized in \( u \) if

\[
a_1 u + u = N [\text{sgn} \, p_2] \tag{29}
\]

where \( N = \max |a_1 u + u| \) for each fixed \( t(0 \leq t \leq T) \). The control \( u^*(t) \) which satisfies these conditions is a “bang-bang” type control where \( u = 1 \) at all times and \( u \) at the moment of switching is unbounded.

Since \( e_3 \) has served its purpose in the optimization process, we may now return to the second order system and solve for the “impulse” variables. By (9),

\[
p = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \tag{30}
\]

and the solution for \( u^*(t) \) becomes

\[
u^*(t) = 1 \cdot \text{sgn} [\cos(t + \Theta)] \tag{31}
\]

where \( \Theta \) is a phase angle dependent on \( x^0 \).

Several properties of the optimum controller are now known. First, the control is a bang-bang type which applies maximum effort at all times in one of the two “directions.” It is switched periodically from one state to the other every half cycle until the origin is reached. Notice that each time the control is switched, a discontinuity appears in the \( e_2 \) variable. This occurs because \( u \) contains an impulse.

\[
\Delta e_2 = \int_{t^-}^{t^+} (-e_1 + a_1 u + u) dt = \int_{t^-}^{t^+} a_1 u dt = a_1 [u(t_1) - u(t_2)]
\]

where \( \Delta e_2 \) is the discontinuity in \( e_2 \) at the time of switching \( t_1 \).

One would now like to find a switching curve \( L(e) \) which divides the phase plane \( e_1 \) vs. \( e_2 \) in such a manner that control \( u^* = +1 \) is optimum in the space to one side of the curve and \( u^* = -1 \) elsewhere. Control would be switched when the trajectory \( e^*(t) \) crosses the curve. The discontinuity \( \Delta e_2 \) precludes this possibility. For example, examine the trajectory \( e^*(t) \) for some initial conditions that dictate \( u^* = -1 \) for optimum control. At the point where this trajectory crosses \( L(e) \) the optimum becomes \( u^* = +1 \). The control switches and \( \Delta e_2 = +2a_1 \) occurs which places the states back in the space where \( u^* = -1 \) was optimum. Here the control switches again, \( \Delta e_2 = -2a_1 \), occurs and chatter motion begins. The fact that \( e_2 \) is multiple valued at the instant of switching makes a simple realization of \( L(e) \) impossible.

For periods between switchings where \( u = 0 \), the system is well behaved with the solution for the \( k^{th} \) interval

\[
e_1(t) = K_1 \cos(t + \phi_1) - \delta
e_1(t) = K_1 \cos(t + \phi_1 + \pi/2) \tag{33}
\]

where \( \delta = 1 \cdot \text{sgn} p_2 \) and \( K, \phi \) depend on conditions of states at the start of the \( k^{th} \) interval.

4.1 The transformed variable

The search for a variable of the system on which to control leads to the possibility of “subtracting out” the discontinuity present in \( e_2 \) at times of switching.

The Laplace transforms of the system variables are

\[
E_1(s) = \frac{e_2 s + e_2 + (a_1 s + 1)U(s)}{s^2 + 1} \tag{34}
\]

\[
E_2(s) = \frac{e_2 s - e_1 + s(a_1 s + 1)U(s)}{s^2 + 1}
\]

where

\[
U(s) = \delta \left( \frac{1}{s} - \frac{2}{s} e^{-t_r} + \frac{2}{s} e^{-t_r} - \ldots \right) \tag{35}
\]

which for any instant of time \( t(0 \leq t < t_1) \)

\[
U(s) = \frac{\delta}{s} \tag{36}
\]
Equations (34) then become

\[ E_1(s) = \frac{e_1^2 s^2 + (e_0^2 + a_1 \delta)s + \delta}{s(s^2 + 1)} \]

\[ E_2(s) = \frac{(e_0^2 + a_1 \delta)s + (-e_0^2 + \delta)}{s^2 + 1} \]

(37)

By means of the initial value theorem, it is seen that

\[ \lim_{t \to 0} e_1(t) = \lim_{s \to \infty} sE_1(s) = e_1^2 \]

\[ \lim_{t \to 0} e_2(t) = \lim_{s \to \infty} sE_2(s) = e_2 + a_1 \delta \]

(38)

At time \( t = 0 \), \( e_2 \) jumps to \( e_2 + a_1 \delta \). To remove this discontinuity consider the transformed variables

\[ Y_1(s) = E_1(s) \]

\[ Y_2(s) = E_2(s) - \frac{a_1 \delta}{s} \]

(39)

By virtue of (39) and (20)

\[ sY_1(s) = sE_1(s) = E_2(s) = Y_2(s) + \frac{a_1 \delta}{s} \]

\[ sY_2(s) = sE_2(s) - a_1 \delta = -Y_1(s) + \frac{\delta}{s} \]

or

\[ y = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \delta + \begin{bmatrix} a_1 \\ 1 \end{bmatrix} \]

(40)

(41)

where \( \delta \) is a unit step function with sign to be determined. The system (41) is identical to that of (20) except for the action at time of switching. It should be noted, however, that care must be taken in assigning final values to the system described by (41) if the two plants are to be controlled in parallel. The final value theorem and (39) gives

\[ \lim_{t \to \infty} Y_2(t) = \lim_{s \to 0} sY_2(s) = \lim_{s \to 0} s \left( E_2(s) - \frac{a_1 \delta}{s} \right) = -a_1 \delta \]

(42)

From this it is observed that zeroing the final states in (20) is analogous to zeroing \( y_1(T) \) and attaining a final value

\[ Y_2(T) = -a_1 \delta \]

in the system (41).

4.2 Boundary conditions and final control

From (38) and (39) it is clear that the initial conditions on the \( e \) and \( y \) variables are identical. From (39) it is also seen that

\[ y_1(T) = e_1(T) \]

\[ y_2(T) = e_2(T) - a_1 \delta(T) \]

(44)

At this point in the pursuit of the optimum control, it becomes necessary to investigate the system action possible at time \( t = T \) under admissible control. \( \Delta e_2 \) of (32) provides a means of changing the value of \( e_2 \) instantaneously by an amount dictated by the constraints on \( u(t) \). With this in mind, it is noted that appropriate use of \( \Delta e_2 \) within the bounds of allowable control may zero the \( e_2 \) variable in zero time given that \( e_2(T) \) is within range. The conditions (21) and (44) with (32) indicate that for

\[ |y_2(T)| \leq a_1 \]

(45)

the system (20) may be zeroed instantly. The boundary conditions on (41) then become

\[ y_i(0) = y_i = e_i \quad i = 1,2 \]

\[ y_1(T) = 0 \]

\[ |y_2(T)| \leq a_1 \]

(46)

The final controller \( u_2(T) \) that must zero the errors for \( t > T \) has two conditions imposed upon it, i.e.,

\[ a_1 u_2 + u_2 = 0 \]

\[ u_2(T) - \delta(T) = \frac{-e_2(T)}{a_1} \]

(47)

The solution to (47) is

\[ u_2(t) = \frac{-y_2(T)}{a_1} \exp \left( -\frac{t + T}{a_1} \right) \quad t \leq T \]

(48)

It is assumed then that \( u_2(t) \) is available at time \( t = T \) so that the boundary conditions on the system are as stated in (46).

4.3 Switching functions

The method of finding a function \( L(y) \) with which to describe the switching criteria for the optimum trajectory proceeds as follows. As in the discontinuous case, it is desired that

\[ S = \int_0^T \alpha dt \]

(49)

* The conditions are stated in terms of the \( y \) variable for convenience in order that notational problems arising from multiple value of \( e_2(0) \) be avoided.
be minimized, therefore, another variable $S = c_yY_2(T)$ is adjoined to the system and once again $c_1 = c_2 = 0$. The Hamiltonian becomes

$$H = P_1Y_2 + P_1a\delta - P_2Y_1 + P_2\delta - \alpha$$

(50)

This is maximized in $\delta$ when

$$\delta = 1 \cdot \text{sgn} (a_1P_1 + p_2)$$

(51)

With this control, trajectories are circular about $(-a_1, 0)$ with radius determined by $y^0$. (See Fig. 3.)

Previous arguments have determined that the conditions on the system are

$$y_i(0) = y_i, \quad i = 1, 2, 3$$

$$y_i(T) = 0$$

$$|y_i(T)| \leq a_1$$

$$H(T) = 0$$

$$p_3(T) = -1$$

(52)

The function

$$F = \frac{1}{2} (y_2 - a_1) \leq 0$$

may be used to describe $G$. From this

$$b_2(y_2(T)) = \left. \frac{\partial F}{\partial y_2} \right|_{y_2 = y_2} = y_2$$

(53)

and

$$p_2(T) = -\lambda c_2 - \mu b_2(y_2(T)) = -\mu y_2$$

(54)

where $\mu \geq 0$ with modulus such that $F(T) = 0$. In the phase plane of $p_1$ vs. $p_2$ it is sufficient to note that for trajectories terminating at $y_1(T) = -a_1$, $p_2(T) \geq 0$ and for trajectories ending at $y_1(T) = a_1, p_2(T) \leq 0$. This information in addition to the control (51) completely define $L(y)$ for trajectories ending on the extremes of the line segment $y_2(T) = a_1$.

Fig. 4 depicts representative action for optimum trajectories terminating at $y_2(T) = -a_1$, $y_1(T) = 0$. Trajectories ending at $y_2(T) = +a_1, y_1(T) = 0$ are mirror images. The optimum switching curves are generated by picking arbitrary values of $p(T)$ from the admissible set for the corresponding boundary values of $y(t)$ and working backwards in time plotting the switching points determined from $p(-t)$ on the $y_1$ vs. $y_2$ phase plane.

For trajectories ending in the interior of the line segment where $|y_2(T)| < a_1, b_2(y_1(T)) = 0$ and, therefore, $p_2(T) = 0$. This completes the information necessary to describe $L(y)$. Fig. 5 shows a representative trajectory arrived at by translating switching criteria from the $p$ plane to the $y$ plane. Fig. 6 portrays the curve with all dimensions.

V. THE MINIMUM FUEL PROBLEM

The minimum fuel problem is solved by minimizing the integral
Figure 5. $p_1$ vs. $p_2$ and $y_1$ vs. $y_2$ Phase Planes with Complete Switching Curves.

\[ T = \int_0^T (|u| + |a_1u|) dt \quad (55) \]

The above formula appears simpler, however, to once again make use of the transformed variable $y(t)$. By minimizing

\[ J = \int_0^T |u| dt \quad (56) \]

in the transformed system (41), the desired result can be obtained provided

i) the switchings in the time interval $0 \leq t \leq T$ are kept to a minimum.

ii) adjustment is made at time $t = T$ when fuel is consumed zeroing the error states $e(t)$ with the exponential control $u_2(T)$.

After adjoining (56) to the system (41), the hamiltonian becomes:

\[ H = p_1y_2 - p_2y_1 + u(a_1p_1 + p_2) - |u| \quad (57) \]

Since $T$ is not specified $H(T) = 0$. With $u(t)$ constrained as before, the control that maximizes $H$ with respect to $p(t)$ is:

\[ u^* = \begin{cases} \text{sgn} (a_1p_1 + p_2) & |a_1p_1 + p_2| \leq 1 \\ 0 & |a_1p_1 + p_2| > 1 \end{cases} \quad (58) \]

5.1 Initial conditions

Taking the time derivative of $H$

\[ \frac{dH}{dt} = (a_1p_1 + p_2) \frac{du}{dt} - \frac{du}{dt} \quad (59) \]

Figure 6. Optimum Switching Curve (minimum time).

Figure 7. Optimum Trajectories of $e(t)$ and $y(t)$ (minimum time).
it can be seen that \( \frac{dH}{dt} = 0 \) if \( \frac{du}{dt} = 0 \). It may also be argued that the change in the hamiltonian with time is zero if

\[
a_1 p_1 + p_2 = \frac{d|u|}{du} = \frac{\Delta u}{\Delta u}
\]

Since \( u(t) \) is switching between \( u = 0 \) and \( u = \pm 1 \) and vice versa, this means that the hamiltonian remains constant if the control is switched at \( a_1 p_1 + p_2 = 1 \cdot \text{sgn} (\Delta u) \). (See Fig. 10.)

By choosing control \( u^*(t) \) the hamiltonian remains at its maximum value, i.e., identically zero from time \( t = 0^+ \) after initial control has been applied until time \( t = T \). This control minimizes the integral (56) but does not necessarily minimize total fuel when fuel consumed at switchings is added. In order to minimize switchings, it appears necessary to choose the degenerate case, i.e., \( u = 0 \) until such time as \( a_1 p_1 + p_2 = 1 \cdot \text{sgn} (\Delta u) \) where \( \Delta u \) is the change in \( u(t) \) when turning the control on. Notice that this choice guarantees that \( H(t) = 0 \) for all \( t, 0 \leq t \leq T \). With this in mind, the problem remains to minimize fuel in the non-degenerate case. For this purpose it will be considered that time \( t = 0 \) is that time when

\[
a_1 p_1 + p_2 = 1 \cdot \text{sgn} (\Delta u)
\]

and initial control is applied.

At \( t = 0 \) it may be verified from (61) and because \( H(0) = 0 \) that

\[
p_1 y_2 - p_2 y_1 = 0
\]
5.2 Final boundary conditions

In order to investigate final value boundary conditions, the optimum trajectories terminating such that \( y_1(T - \Delta t) > 0 \) are considered. Trajectories in the rest of the space are mirror images. As in the minimum time problem, an optimum trajectory terminating at \( y_2(T) = -a_1, y_1(T) = 0 \) is investigated first. The determination that \( p_2(T) \geq 0 \) as argued in (54) is still valid. This condition on \( p_2(T) \) along with the fact that \( H(T) = 0 \) precludes the possibility of a trajectory terminating as above with \( u(T) = -1 \). The following cases, however, do apply. Consider

\[
H(T) = -a_1p_1(T) + u(T) [a_1p_1(T) + p_2(T)] - |u(T)| = 0
\]

This condition implies that if \( u(T) = 0 \) then \( p_1(T) = 0 \) and if \( u(T) = +1 \) then \( p(T) \geq 0 \) and \( p_2(T) = +1 \). Fig. 11 portrays the locus of admissible points \( p(T) \) and the switching curves generated by these criteria in the \( y_1 \) vs. \( y_2 \) phase plane as in Fig. 11a.

Optimum trajectories terminating on the line segment \( y_1(T) = 0, |y_2(T)| < a_1 \) must be investigated in a fashion similar to that used with the minimum time problem. Since a final boundary point \( y_2(T) \) is not fixed, we may substitute a final condition on \( p_2(T) \) to reach a solution. At this point it becomes necessary to decide on the final value functional to be minimized. It is first noted that if the final control to the line segment is \( u(T) = 0 \), then \(-a_1 < y_2(T) < 0\). (It must be remembered that investigation is of trajectories such that \( y_1(T - \Delta t) > 0 \)). If \( u(T) = 0 \) then also \( y_2(T) = e_2(T) \) and in order to minimize the fuel consumed by \( u_2(T) \) to zero \( e_2 \) after time \( T \) then \( |e_2(T)| = |y_2(T)| \) must be minimized.

If the final control is \( u(T) = -1 \) (\( u(T) = +1 \) is not possible for trajectories terminating on this side of the line segment) then \( e_2(T) = y_2(T) + a_1 \) and, therefore, \(|y_2(T) - a_1| \) must be minimized. In both of the above cases, it may be seen that \( y_2(T) \) must be maximized on the line segment in order that fuel consumed by \( u_2(T) \) to zero the error states be minimized. Therefore, the functional to be minimized is

\[
S = \sum_{i}^{3} c_i y_i(T) = -y_2(T) + y_3(T) \tag{64}
\]

where

\[
y_3(t) = \int_{0}^{t} |u| dt \tag{65}
\]

By prior arguments \( p_2(T) = -c_2 = +1 \) and \( p_3(T) = p_3(t) = -c_3 = -1 \).

5.3 Generating the switching curve segments

It is now helpful to look at the hamiltonian under each of the above conditions, i.e., \( u(T) = 0 \) and \( u(T) = -1 \). In the first case

\[
\begin{align*}
u(T) &= 0 \\
H(T) &= H(t) = 0 \\
y_1(T) &= 0 \\
p_2(T) &= +1 \\
-a_1 &< y_2(T) < 0
\end{align*}
\]

Figure 11a. Switching Criteria for \( y_2(T) = -a_1 \) (minimum fuel).
and
\[
H(T) \ p_1(T) y_z(T) = 0
\]
which implies that \( p_1(T) = 0 \). Fig. 12 shows the switching generated by this condition.

![Figure 12. Switching Criteria for \( u(T) = 0 \) (minimum fuel).](image)

Next is considered the case where
\[
u(T) = -1
\]
\[
H(T) = H(t) = 0
\]
\[
y_1(T) = 0
\]
\[
p_2(T) = +1
\]
and
\[
H(T) = p_1(T) y_z(T) - 1 \left[a_1 p_1(T) + 1\right] - 1 = 0
\]
from which
\[
p_1(T) = \frac{2}{y_z(T) - a_1}
\]
(68)

Since \( u(T) = -1 \) and \( p_2(T) = +1 \), conditions (58) are met only when
\[
p_1(T) < \frac{-2}{a_1}
\]
which implies \( y_z(T) > 0 \). In Fig. 13 these trajectories and switching curves are plotted.

![Figure 13. Admissible \( p(T) \) where \( u(T) = -1 \).](image)

5.4 The complete switching curve

Because \( T \) was never specified and because the fuel consumed at switching was handled as a side condition, a composite of all the calculated switching curves indicates areas in the phase plane where criteria for optimum control appear contradictory. In these areas, analysis by graphical means or actual computation will clear up the situation. Fig. 14 depicts the composite of the first two criteria analyzed.

![Figure 14. Region of Conflicting Optimum Criteria.](image)

In Fig. 14, region A is an area where there is a question concerning whether it is optimum to switch for \( |y_z(T)| = a_1 \) or \( |y_z(T)| < a_1 \). By graphical analysis, it may be seen that it is optimum to switch so that \( |y_z(T)| = a_1 \).

A similar contradiction between trajectories switching for \( 0 < y_z(T) < a_1 \) and \( -a_1 < y_z(T) < 0 \) may also be resolved graphically.* The final

* Appendix I presents computational analysis of the resolving process.
result consisting of switching criteria to zero the errors in the system (20) with minimum fuel is given by Fig. 15.

![Figure 15. Switching Criteria for Minimum Fuel, \( a_1 = 1.0 \).](image)

**VI CONCLUSIONS**

The methods used in this paper to arrive at a solution may be used to good advantage in the investigation of any \( n \)th order system with no more than \( n-1 \) zeros. The maximum principle provides a powerful tool in optimization, particularly for linear systems. Often the method of Pontryagin will indicate areas of interest to investigate when searching for an optimum control even if the unique solution is not readily forthcoming.

The problem of controlling a plant with zeros is analogous to controlling a plant without zeros using an impulse-step type controller. Results obtained in this paper can be adapted to formulate the logic of this type control.

The realization of the true optimum switching logic in a practical system may in many cases not be worth the effort. Quasi-optimum control using simple switching functions that are for the most part linear is a subject for further investigation. Setting time for the system is relatively insensitive to limited variations from the optimum when trajectories are out beyond the first cusp of the switching curve.

**APPENDIX I**

In the past, the electrical engineer, has usually turned to the analog computer for problem solving and his reference of familiarity is strongest there. To test the drift of the ampli-

![Figure A-1. Trajectories showing the effects of different integration step sizes (\( \Delta \)) on the Runge-Kutta method. Period of unforced system being integrated was \( T = 6.2836 \).](image)
DISCONTINUOUS SYSTEM VARIABLES IN THE OPTIMUM CONTROL OF SECOND ORDER

A common method was to plug in a simple oscillatory system

\[ y = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} y \]

and to study the decay of the circular trajectory in the phase plane (y vs. y).

This same procedure was used here to study the different integration schemes used and to determine the proper integration step size. The forcing function \((u = \pm 1)\) was introduced to give breaks in slopes and also discontinuities in the steady state limit cycle described by

\[ y = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} y - \begin{bmatrix} 1 \\ 1 \end{bmatrix} \text{sgn} (y_1 - y_2) \]

This graphical means of analyzing gave considerable insight in evaluating the integration schemes. The following figures give a brief résumé of these studies.

The 4th order Runge-Kutta method was found to be the most foolproof when dealing with discontinuities in the state variables. In the continuous variables the Adams Bashforth predictor-corrector with a Taylor expansion or Runge-Kutta starter were found efficient and useful.

The system was also tested with the plant on the analog computer and the control with its nonlinear switching function being generated on an on-line digital computer.

BIBLIOGRAPHY


