Effects of Propellant Mass Loss on Fuel-Optimal Rendezvous Near Keplerian Orbit

Thomas E. Carter*
Eastern Connecticut State University, Willimantic, Connecticut

Previous work on fuel-optimal rendezvous of a spacecraft near a point in arbitrary Keplerian orbit has been generalized by consideration of the change in the mass of the spacecraft as the propellent is depleted. Although the Newtonian gravitational force function has been linearized in describing the relative motion of the spacecraft, the total system of equations is nonlinear because the mass is not constant in this formulation. Bounded thrust, fixed time, and constant exhaust velocity are assumed. Necessary conditions for solution of this problem are presented that reveal the structure of an optimal rendezvous. Most of the results are based on an examination of the switching function. It is found that inclusion of the mass-variation effect causes very little change in the nature of the singular solutions. The nature of the nonsingular solutions, however, is altered by a change in the location of the switches between thrusting and coasting intervals. These switches are determined by a dynamic geometric relationship between a primer vector and an expanding sphere. If the mass is constant, a more static geometric relationship develops.

I. Introduction

MANY fuel-optimal trajectory or rendezvous problems can be simplified by consideration of the relative motion of a spacecraft with respect to a real or imaginary object in a Keplerian orbit. Possible examples include maneuvers of a spacecraft near an orbiting space station, or a small asteroid or comet. Other types of examples include correcting the accumulated position and velocity errors from general space missions by attempting to rendezvous with a fictitious object in the correct Keplerian orbit. These types of problems can be approximated by linearizing a gravitational force function about a point in Keplerian orbit.

This approach was taken recently[1] to define the problem of fixed-duration, bounded-thrust, fuel-optimal rendezvous of a spacecraft near a point in general Keplerian orbit, and its solution was investigated assuming constant spacecraft mass. We extend that investigation here by including the effect of the change in the mass of the spacecraft as the propellant is burned. This also generalizes other fixed-mass studies in which the orbit is circular.[2,3] A comparison of these new results with those previous ones reveals many similarities and some significant differences.

The transformed, linearized equations of motion that we use were derived for a spacecraft near a satellite in an elliptical orbit by De Vries,[4] Tschauner and Hempel,[5] Shulman and Scott,[6] and, for the planar case, Euler.[7] The homogeneous form of these equations was found and they were solved by Lawden[8] much earlier in an entirely different context to describe his "primer vector." The rendezvous problems investigated in Refs. 5 and 9 are similar to the recent investigation[1] that we generalize here. More details, other references, and another derivation of the equations can be found in that study. With similar notation used, the equations that immediately follow are taken from Ref. 1.

All vectors will be assumed to be elements of three-dimensional Euclidean space. The independent variable θ, the true anomaly, is the angle of the orbiting body measured from periapsis. The position vector $\boldsymbol{x}(\boldsymbol{\theta})$ of the spacecraft is measured from the center of the orbiting object in a rotating coordinate frame, in which the positive x_2 axis is directed away from the center of attraction, the x_1 axis is perpendicular to it in the plane of the orbit and its positive direction opposes the motion of the orbiting body, and the x_3 axis completes a right-handed system. The position vector is transformed to a new variable $y(\theta)$ by the transformation

$$y(\theta) = (1 + e\cos\theta)x(\theta) \tag{1}$$

where $e \geq 0$ denotes the eccentricity of the orbit. The vector $y(\theta)$ is proportional to $x(\theta)$ divided by the distance of the orbiting body from the center of attraction, and satisfies a more concise set of equations of motion than $x(\theta)$. If we use subscripts to denote the components of y, these linear differential equations are

$$y_1''(\theta) = 2y_2'(\theta) + a_1(\theta) \tag{2a}$$

$$y_2''(\theta) = \frac{3y_2(\theta)}{1 + e\cos\theta} - 2y_1'(\theta) + a_2(\theta) \tag{2b}$$

$$y_3''(\theta) = -y_3(\theta) + a_3(\theta) \tag{2c}$$

where the prime indicates differentiation with respect to θ and the transformed acceleration vector $a(\theta)$ is given by

$$a(\theta) = b\frac{u(\theta)/m(\theta)}{(1 + e\cos\theta)^3} \tag{3}$$

In this expression, $u(\theta)$ denotes a normalized thrust vector satisfying the condition

$$|u(\theta)| \leq 1 \tag{4}$$

where the notation $||$ is used to indicate the magnitude or Euclidean norm of a vector. The positive constant b is L^6T_m/μ^4 where μ is the universal gravitational constant times the mass of the central body of attraction, L the magnitude of the constant angular momentum of the orbiting object divided by its mass, and T_m the maximum magnitude of the thrust of the spacecraft. Note that this definition of b is slightly different from that of the previous work[1] in that it does not include the spacecraft mass. If we assume that the exhaust velocity c of the propellant is constant, the mass $m(\theta)$ of the spacecraft satisfies

Received May 29, 1987; revision received Sept. 25, 1987.

*Professor, Department of Mathematics and Computer Science.

the equation

$$m'(\theta) = \frac{-k|u(\theta)|}{(1+e\cos\theta)^2} \tag{5}$$

where the constant k is $\mu^2 b/(L^3 c)$. Although Eqs. (2) are linear in y, we see that the total system is not linear because of the inclusion of the variable mass in Eqs. (3) and (5).

II. Necessary Conditions for Fuel-Optimal Rendezvous

We shall assume that the initial and final values θ_0 and θ_f of θ are fixed and so formulate a fixed-duration rendezvous problem. We let Θ denote the fixed interval $\theta_0 \le \theta \le \theta_f$ and note that since the orbit is Keplerian, then $-\cos^{-1}(-1/e) < \theta_0 < \theta_f < \cos^{-1}(-1/e)$ if $e \ge 1$.[1] We define the class of admissible control functions as the set of all Lebesgue measurable vector-valued functions that satisfy Eq. (4) a.e. on Θ. We shall assume that the preceding equations (2), (3) and (5) are defined a.e. on Θ, and when placed in state vector form, become

$$y'(\theta) = v(\theta) \tag{6a}$$

$$v'(\theta) = A(\theta)y(\theta) + Bv(\theta) + \frac{b}{(1+e\cos\theta)^3}\frac{u(\theta)}{m(\theta)} \tag{6b}$$

$$m'(\theta) = -\frac{k|u(\theta)|}{(1+e\cos\theta)^2} \tag{6c}$$

where the matrices $A(\theta)$ and B are given by

$$A(\theta) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 3/1+e\cos\theta & 0 \\ 0 & 0 & -1 \end{pmatrix} \qquad B = \begin{pmatrix} 0 & 2 & 0 \\ -2 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \tag{7}$$

The positive number m_0 represents the initial mass of the spacecraft, the arbitrary vectors y_0 and v_0 represent the initial transformed position of the spacecraft and its derivative where the transformation is given by Eq. (1), and similarly, the arbitrary vectors y_f and v_f denote the final transformed position and velocity. We formulate the fuel-optimal rendezvous problem as follows.

We seek an admissible control function u to minimize the integral

$$J[u] = \int_{\theta_0}^{\theta_f} \frac{|u(\theta)|\,d\theta}{(1+e\cos\theta)^2} \tag{8}$$

subject to the differential equations (6), the initial conditions

$$y(\theta_0) = y_0, \qquad v(\theta_0) = v_0, \qquad m(\theta_0) = m_0 \tag{9}$$

and the terminal conditions

$$y(\theta_f) = y_f, \qquad v(\theta_f) = v_f \tag{10}$$

An alternate formulation of this problem is to find an admissible control function u to maximize $m(\theta_f)$ subject to the differential equations (6) and the end conditions (9) and (10). This formulation does not directly involve the integral (8). The reason that we do not use this second formulation is that it is not general enough to include the case where $k = 0$ in Eqs. (6). Our formulation enables us to consider also the case in which the spacecraft mass is constant simply by setting $k = 0$ in our results.

Application of Control Theory

The well-known principle of Pontryagin provides necessary conditions for a minimum fuel solution of this rendezvous problem. For the case in which the mass is constant ($k = 0$), it also provides sufficient conditions because the integrand in Eq. (8) is independent of the state and the differential equations are linear (Ref. 10, Sec. 5.2.). Since our Eqs. (6) are nonlinear for this variable mass problem, we cannot assert sufficiency conditions. For this reason, an admissible control u and an associated solution of the differential equations (6) that satisfy Pontryagin's principle will be referred to as *extremal* instead of optimal. Clearly, an optimal solution is also extremal.

Using the principle of Pontryagin, we can assert that if u is an admissible control such that (y,v,m,u) is a solution of the fuel-optimal rendezvous problem, then a.e. on Θ $u(\theta)$ minimizes the pre-Hamiltonian

$$H(u_0) = \frac{l_0|u_0|}{(1+e\cos\theta)^2} + p(\theta)^T v(\theta) + q(\theta)^T\left[A(\theta)y(\theta) + Bv(\theta) + \frac{bu_0}{(1+e\cos\theta)^3 m(\theta)}\right] - \frac{l(\theta)k|u_0|}{(1+e\cos\theta)^2} \tag{11}$$

over the set $|u_0| \le 1$ where the notation is abused for brevity; the superscript T is used to indicate the transpose of a matrix or vector, $l_0 \ge 0$; and the vector functions p and q and the scalar function l are adjoint variables that a.e. on Θ satisfy the differential equations

$$p'(\theta) = -A(\theta)^T q(\theta) \tag{12a}$$

$$q'(\theta) = -p(\theta) - B^T q(\theta) \tag{12b}$$

$$l'(\theta) = \frac{bq(\theta)^T u(\theta)}{(1+e\cos\theta)^3 m(\theta)^2} \tag{12c}$$

Problems in which $l_0 = 0$ are called *abnormal*; otherwise, they are called *normal*. We can normalize by dividing Eq. (11) by l_0 so that, effectively, l_0 can be set equal to any positive number. In order to include both normal and abnormal problems, we shall not yet specify a value of l_0.

We see that $u(\theta)$ minimizes Eq. (11) if it minimizes the expression

$$L(u_0) = [l_0 - kl(\theta)]|u_0| + \frac{bq(\theta)^T u_0}{(1+e\cos\theta)m(\theta)} \tag{13}$$

where again for brevity, the notation is abused. We may write $u(\theta) = e_u(\theta)f(\theta)$ where $e_u(\theta)$ is a unit vector and

$$0 \le f(\theta) \le 1 \tag{14}$$

a.e. on Θ because of the bound [Eq. (4)]. Making this kind of substitution in Eq. (13), we observe that if $q(\theta) \ne 0$, we can minimize $L(u_0)$ by setting u_0 equal to $e_u(\theta)f(\theta)$, where $e_u(\theta) = -q(\theta)/|q(\theta)|$, and by picking $f(\theta)$ to minimize Eq. (13) that is now in the form $s(\theta)f(\theta)$, where the function s is defined by

$$s(\theta) = l_0 - kl(\theta) - \frac{b|q(\theta)|}{(1+e\cos\theta)m(\theta)} \tag{15}$$

This function s will be called the *switching function*. It follows that unless $|q(\theta)|$ is zero on a set of positive measure, then an optimal control function must be given by

$$u(\theta) = -\frac{q(\theta)}{|q(\theta)|}f(\theta) \tag{16}$$

a.e. on Θ where

$$f(\theta) = \begin{cases} 0, & s(\theta) > 0 \\ 1, & s(\theta) < 0 \end{cases} \tag{17}$$

or if $s(\theta) = 0$, $f(\theta)$ takes any value that satisfies Eq. (14). In case the zeros of q (i.e., roots of $|q(\theta)| = 0$) form a set of positive measure, then it follows from the fact that the components of

q are analytic functions[1] of a real variable with infinitely many zeros on a bounded interval that q is identically zero; therefore, Eqs. (12) imply that $l(\theta)$ is a constant. From the transversality condition $l(\theta_f) = 0$, we get $l(\theta) = 0$ for each $\theta \in \Theta$. Since optimal control theory does not allow p, q, l, and l_0 to all be zero at θ_f, we must have $l_0 > 0$. The problem must be normal, and Eq. (13) can be minimized only by $|u(\theta)| = 0$. This shows that if $|q(\theta)|$ is zero on a set of positive measure, then it is zero identically and we must have $u(\theta) = 0$ a.e. on Θ. This extremal control and related trajectory we call the trivial extremal. It consists of coasting without thrusting on the entire interval Θ. Except for a trivial solution such as this, an optimal control function must satisfy Eq. (16) a.e. on Θ, where f satisfies Eq. (17) if $s(\theta) \neq 0$ and inequality (14) otherwise. Having discussed the trivial extremal, we shall now assume that $q(\theta)$ is nonzero a.e. on Θ.

If the switching function s is nonzero on a set S of positive measure, then it is necessary that an optimal control function satisfies Eqs. (16) and (17) a.e. on S. An optimal or extremal control function and its associated trajectory will therefore be called *nonsingular* on S if $s(\theta) \neq 0$ a.e. on S. On the other hand, if $s(\theta) = 0$ a.e. on a set S of positive measure, it is necessary that an optimal control satisfies Eq. (16) and inequality (14) a.e. on S, and any control function that satisfies inequality (14) and Eq. (16) minimizes Eq. (13) and is therefore extremal. We therefore call an optimal or extremal control function and its associated trajectory *singular* on S if $s(\theta) = 0$ a.e. on S.

Robbins[11] has shown that for certain nonlinear fuel-optimal problems, an extremal control may be comprised of both singular and nonsingular subarcs. For this reason, we define singular and nonsingular solutions on a set S of positive measure, rather than on the entire set Θ. Later, we will answer the question of whether or not singular and nonsingular regimes, each of positive measure, can coexist on an extremal.

The key to further understanding of either the singular or nonsingular extremals for this problem lies in an examination of the switching function [Eq. (15)] that depends on the adjoint variables $q(\theta)$ and $l(\theta)$ defined by Eqs. (12). Eliminating $p(\theta)$ and noting from Eqs. (7) that $A(\theta)^T = A(\theta)$ and $B^T = -B$, we observe that $q(\theta)$ is defined by the differential equation

$$q''(\theta) - Bq'(\theta) - A(\theta)q(\theta) = 0 \tag{18}$$

Following Lawden,[12] we will refer to q as the primer vector. We note from Eqs. (6) that y is defined by the differential equation

$$y''(\theta) - By'(\theta) - A(\theta)y(\theta) = \frac{b}{(1 + e\cos\theta)^3}\frac{u(\theta)}{m(\theta)} \tag{19}$$

We see from Eqs. (16) and (17) that on intervals where $s(\theta) > 0$ (coast intervals), the differential equations (18) and (19) are identical. This phenomenon was observed in earlier work[1,2] in which the problems were linear. As previously found,[1] it is convenient to refer to the transformed primer Q defined by

$$Q(\theta) = \frac{q(\theta)}{1 + e\cos\theta} \tag{20}$$

In terms of the transformed primer, an extremal control function [Eq. (16)] becomes

$$u(\theta) = -\frac{Q(\theta)}{|Q(\theta)|}f(\theta) \tag{21}$$

a.e. on Θ and the switching function (15) becomes

$$s(\theta) = l_0 - kl(\theta) - b|Q(\theta)|/m(\theta) \tag{22}$$

where the adjoint variable $l(\theta)$ that satisfies Eqs. (12) now satisfies

$$Z'(\theta) = -\frac{b|Q(\theta)|f(\theta)}{(1 + e\cos\theta)^2 m(\theta)^2} \tag{23}$$

The differential equation (18) can be solved so that the transformed primer Q can be found through Eq. (20). This solution also appeared in previous work[1] in slightly different form for the case $e \neq 0$. We shall treat separately the case where $e = 0$. Using subscripts, we write components as

$$Q_1(\theta) = -\frac{b_1}{e}r(\theta) - \frac{b_2}{e}\left[r(\theta)I(\theta) + \frac{\cos\theta}{r(\theta)\sin\theta}\right] - \frac{c_1}{e}\sin\theta\left[1 + \frac{1}{r(\theta)}\right] + \frac{c_2}{r(\theta)} \tag{24a}$$

$$Q_2(\theta) = [b_1 + b_2 I(\theta)]\sin\theta - \frac{c_1}{e}\cos\theta \tag{24b}$$

$$Q_3(\theta) = \frac{\alpha\sin\theta + \beta\cos\theta}{r(\theta)} \tag{24c}$$

where

$$r(\theta) = 1 + e\cos\theta \tag{25}$$

and

$$I(\theta) = \int\frac{d\theta}{\sin^2\theta\, r(\theta)^2} \tag{26}$$

and b_1, b_2, c_1, c_2, α, and β are arbitrary constants of integration. This integral has been evaluated[1,8] and the particular form is dependent on the value of e. (The evaluation in Ref. 1 for the case $e = 1$ is in error). Only if $e = 0$ or $e = 1$ is it a ratio of trigonometric polynomials. Although the components of Q may not appear to be analytic functions, the singularities that occur where $\sin\theta$ is zero are removable in Q and its derivatives so that with this adjustment, they are analytic on Θ.[1] We shall make use of this fact later.

Two Expressions Involving the Switching Function

Additional information about the nature of an optimal solution can be found from an examination of the switching function [Eq. (22)]. We shall use the switching function in the following two ways.

First, we differentiate Eq. (22) to obtain

$$s'(\theta) = -kl'(\theta) + b|Q(\theta)|m'(\theta)/m(\theta)^2 - b|Q(\theta)|'/m(\theta) \tag{27}$$

If we replace $l'(\theta)$ and $m'(\theta)$ using Eqs. (23) and (5), where $|u(\theta)| = f(\theta)$, Eq. (27) becomes the remarkably simple expression

$$s'(\theta) = -b|Q(\theta)|'/m(\theta) \tag{28}$$

This condition will be especially useful in investigating singular solutions.

Second, we put Eq. (22) into a form where it is more convenient to make observations. We multiply both sides of Eq. (22) by $m(\theta)$ and define ξ by

$$\xi(\theta) = m(\theta)[l_0 - kl(\theta)] \tag{29}$$

Equation (22) can then be written as

$$m(\theta)s(\theta) + b|Q(\theta)| = \xi(\theta) \tag{30}$$

Differentiating Eq. (29) and again using Eqs. (23) and (5), we find that ξ can be written without l as follows:

$$\xi(\theta) = -k\int_{\theta_0}^{\theta} \frac{s(\zeta)f(\zeta)\,d\zeta}{(1+e\cos\zeta)^2} + C \tag{31}$$

where C is a constant of integration, which through Eq. (29) we see is $m(\theta_0)[l_0 - kl(\theta_0)]$. From Eqs. (30) and (31), we obtain the second expression

$$m(\theta)s(\theta) + k\int_{\theta_0}^{\theta} \frac{s(\zeta)f(\zeta)\,d\zeta}{(1+e\cos\zeta)^2} = C - b|Q(\theta)| \tag{32}$$

This condition provides information about nonsingular extremals at points where s is zero and changes sign called switches. One observes from Eq. (17) that a switch denotes an instantaneous transition between thrusting and coasting. In recent studies,[1,2] we found that the switches were determined from intersections of the transformed primer with the unit sphere. We can observe from Eq. (32) that this geometric condition does not hold for problems in which the mass of the spacecraft is not constant. A more general geometric condition does hold, however, which we shall demonstrate after stating some properties of the function ξ.

Because the switching function s is continuous, we note from Eq. (17) that the measurable sets on which an extremal is nonsingular must be comprised of relatively open intervals of full thrust and coast separated by isolated switches or by singular regimes, and we cannot immediately reject more general measurable sets for the singular regimes. We recall that an extremal control is singular on a set of positive measure if and only if the switching function is zero a.e. on the set. From these facts and from Eqs. (17), (30), and (31), we establish the following properties of the function ξ:

1) ξ is monotone nondecreasing.
2) ξ is strictly increasing on a nonsingular full-thrust interval.
3) ξ is constant on a nonsingular coast interval.
4) ξ is positive at the switches.
5) ξ' is zero on a singular regime S and ξ is a positive constant if S is an interval.

The monotone growth of ξ is a result of the inclusion of the variable mass of the spacecraft in the problem formulation. We observe that if we assume the mass is constant, then $k = 0$ and Eq. (31) shows that $\xi(\theta)$ is the constant C. In this case, Eq. (30) shows that the switches can only occur at points where the transformed primer Q intersects the sphere of radius C/b. For normal problems, we can pick l_0 in Eq. (29) so that $C = b$ and the sphere has unit radius.

With the inclusion of variable mass, a more dynamic geometric condition develops. We see from Eq. (30) that at a switch $\theta_s \in \Theta$, we have $b|Q(\theta_s)| = \xi(\theta_s)$, so that the transformed primer Q intersects a sphere of radius $\xi(\theta_s)/b$. We visualize a growing sphere of radius $\xi(\theta)/b$. At the switches, the transformed primer Q intersects this expanding sphere. Moreover, the sphere is constant during coasting intervals where Q is inside the sphere, but it grows at a positive rate during thrusting intervals where Q is outside the sphere. The singular solutions occur on sets of positive measure where Q is on the surface of the sphere.

Special Case of Circular Orbit

Because most of the immediate space missions involve an object in circular orbit, the case where $e = 0$ deserves special attention.

In this case, Eq. (20) shows that $Q(\theta) = q(\theta)$, and Eqs. (24–26) should be replaced by the equations defining Lawden's primer for circular orbit[2] with the form of the equations slightly altered to fit our development here

$$q_1(\theta) = 2\rho\sin\theta + 3C_1\theta + C_2 \tag{33a}$$

$$q_2(\theta) = \rho\cos\theta + 2C_1 \tag{33b}$$

$$q_3(\theta) = \alpha\sin\theta + \beta\cos\theta \tag{33c}$$

where $\rho \geq 0$, C_1, C_2, α, and β are arbitrary constants of integration and, since the orbit is circular with constant angular velocity [Ref. 1, Eq. (9),

$$\theta = \frac{\mu^2}{L^3}(t - t_0) + \theta_0 \tag{34}$$

where t represents the instantaneous time and t_0 the initial time.

The switching function Eq. (22) remains the same, except that q replaces Q, but Eq. (23) becomes

$$l'(\theta) = -b|q(\theta)|f(\theta)/m(\theta)^2 \tag{35}$$

The first condition that must be satisfied by the switching function becomes

$$s'(\theta) = -b|q(\theta)|'/m(\theta) \tag{36}$$

which is easier to use than Eq. (28) because the expression for the primer for circular orbit Eq. (33) is simpler than the expression for the transformed primer [Eqs. (24–26)] that holds for $e > 0$. The second condition is obtained by replacing Eq. (30) by

$$m(\theta)s(\theta) + b|q(\theta)| = \xi(\theta) \tag{37}$$

and Eq. (31) by

$$\xi(\theta) = -k\int_{\theta_0}^{\theta} s(\zeta)f(\zeta)\,d\zeta + C \tag{38}$$

The use of the second condition is presented graphically in Figs. 1–4 to compare nonsingular fuel-optimal thrusting and coasting sequences assuming constant spacecraft mass ($k = 0$) with similar necessary thrusting and coasting sequences assuming variable mass ($k > 0$). Since we are using the same primer function in both cases, we observe that the end conditions $y(\theta_f)$, $v(\theta_f)$, and $m(\theta_f)$ are expected to be different. Because of Eq. (34), we may replace the true anomaly θ by time in Figs. 2 and 4. It is known that Lawden's primer can take a cycloidal shape in the planar case ($\alpha = \beta = 0$), and the constants ρ, C_1, C_2, and θ_0 can be chosen so that it intersects the unit circle at six points.[2] This situation is depicted in Fig. 1, and the magnitude of this primer is shown in Fig. 2. Because we are assuming constant mass in these two figures, they are obtained from Eqs.

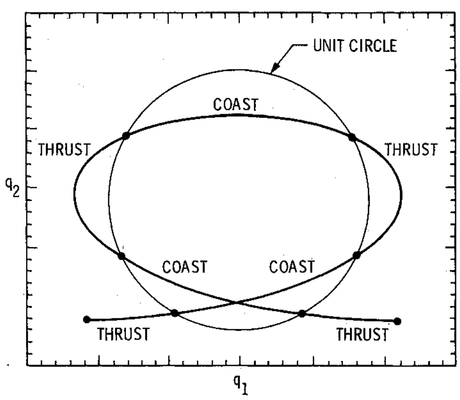

Fig. 1 Fuel-optimal thrusting sequence determined by Lawden primer and unit circle for spacecraft of constant mass.

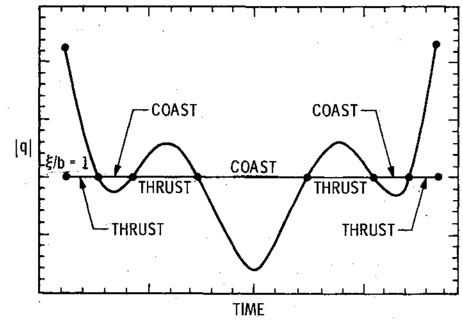

Fig. 2 Magnitude of Lawden primer for spacecraft of constant mass.

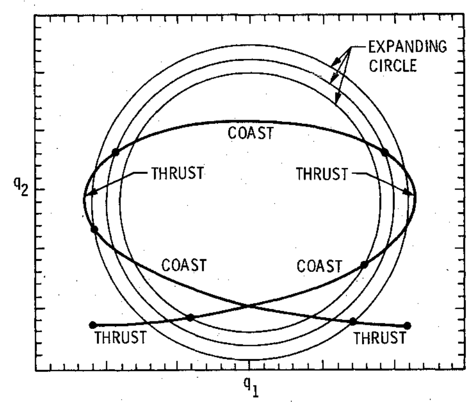

Fig. 3 Fuel-optimal thrusting sequence determined by Lawden primer and expanding circle for spacecraft of variable mass.

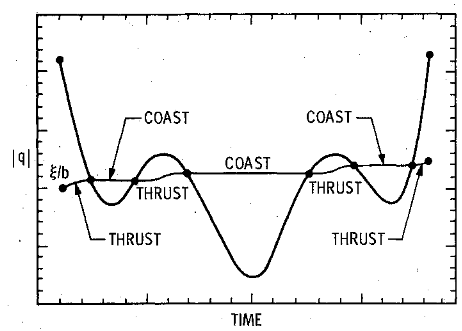

Fig. 4 Magnitude of Lawden primer and ξ/b function for spacecraft of variable mass.

(37) and (38) with $k = 0$ and $C/b = 1$. In Figs. 3 and 4, we remove the assumption of constant mass so that $K > 0$, but use the same Lawden primer and again set $C/b = 1$. We see in Fig. 3 that at the switching points θ_i $(i = 1, \ldots, 6)$, the primer intersects a circle of radius $\xi(\theta_i)/b$ $(i = 1, \ldots, 6)$ according to Eqs. (37) and (38), and we recall the properties of the function ξ previously listed. Figure 4 demonstrates the nature of the function ξ/b that is strictly increasing on thrusting intervals and constant on coasting intervals and that intersects the magnitude of the Lawden primer at the switching times.

III. Investigation of the Zeros of the Switching Function

Much insight into the structure of the extremals is obtained from an examination of the switching function Eq. (22). From this function, we have already established the expressions (28) and (32). We are especially interested in the roots of the equation $s(\theta) = 0$, i.e., the zeros of the switching function. If there are finitely many zeros of this function, then clearly the number of switches is less than or equal to the number of zeros and the location of each switch is defined by the location of a zero. If there are infinitely many zeros, we shall show that the extremal must be singular. For completeness, it is important to determine under what conditions an extremal is singular, as well as to obtain information about the nature of the singular extremal controls, if they exist. This investigation falls naturally into two cases, the noncircular Keplerian orbits ($e > 0$) and the circular orbits ($e = 0$).

Number of Switches

We shall establish the following theorem that shows that an extremal control can have, at most, finitely many switches.

Theorem: The switching function Eq. (22) either has finitely many zeros on a closed bounded interval Θ or else it is identically zero.

Proof: Suppose the switching function (22) has infinitely many zeros on Θ. It follows from Rolle's Theorem that its derivative (28) also has infinitely many zeros on Θ; consequently, the equation

$$|Q(\theta)|' = 0 \tag{39}$$

has infinitely many roots on Θ. It follows therefore from the definition of the Euclidean norm that the equation

$$Q_1(\theta)Q_1'(\theta) + Q_2(\theta)Q_2'(\theta) + Q_3(\theta)Q_3'(\theta) = 0 \tag{40}$$

also has infinitely many roots on Θ. As we have indicated previously, the components of Q are analytic functions of a real variable; consequently, the left-hand side of Eq. (40) is analytic. It is known that an analytic function with infinitely many zeros on a closed bounded interval is zero everywhere so Eq. (40) holds everywhere on Θ. For this reason Eq. (39) is true identically on Θ, and Eq. (28) therefore shows that the switching function s is constant. Since this function has zeros, the constant must be zero, and the proof is complete.

In other words, the aforementioned theorem says the following: *An extremal control either has finitely many switches or it is singular.*

Although we have established that a nonsingular extremal control has finitely many switches, the exact number is not known. Algebraically, the number of switches is limited by the number of zeros of the switching function. Geometrically, it is limited by the number of intersections of the locus of the transformed primer Q, with the expanding sphere whose radius is given by $\xi(\theta)/b$. We conjecture that the number of switches is not more than six if $e \geq 1$ or $\theta_f - \theta_0 \leq 2\pi$. It has been demonstrated[2] that six switches can occur for the case of circular orbit ($e = 0$), constant mass, and $\theta_f - \theta_0 = 2\pi$. The conjecture that no more than six switches can occur, even for this special case, is an unsolved problem.[13]

Singular Extremals

We shall show here that the results that have previously been established[1,3] about singular solutions under the assumption of the constant mass generalize also to this problem ($k \neq 0$).

The first result is a corollary of the recently established theorem because the switching function is identically zero if it is zero at infinitely many points on Θ.

Corollary: If an extremal is singular on any measurable subset of Θ having positive measure, it is singular on Θ. Similarly, it is nonsingular on Θ if it is nonsingular on any measurable subset of positive measure.

This fact is significant because Robbins[11] has shown that a composite extremal of singular and nonsingular subarcs can exist for the original problem without linearization about a point in orbit. Although we have linearized about a point in orbit, our Eqs. (6) are nonlinear. We now state our second result.

Theorem: There are no singular extremals if $e > 0$.

Proof: We assume that $e > 0$ and that an extremal is singular on a set of positive measure. From the preceding corollary the extremal is singular entirely on the interval Θ and the switching function Eq. (22) is identically zero. It follows from Eq. (28) that Eq. (39) holds identically and we see that $|Q(\theta)|$ is constant on θ.

We shall now evaluate this constant. Evaluating the switching function (22) at the end point θ_f and using the transversality condition $l(\theta_f) = 0$, we find that this constant is $l_0 m(\theta_f)/b$. Since control theory does not allow all the adjoint variables to be zero at θ_f and $l_0 = 0$ also, we see that if an extremal is singular, it cannot also be abnormal. A singular extremal is therefore normal and we may pick l_0 to be any positive number. For convenience, we now set

$$l_0 = b/m(\theta_f) \tag{41}$$

and with this definition of l_0, we have established that an extremal is singular if and only if

$$|Q(\theta)| = 1 \tag{42}$$

for each $\theta \in \Theta$. From this point on, arguments that were used in an earlier study[1] in which constant mass was assumed also apply here. Using those arguments, we see that for $e > 0$, Eq. (42) cannot be satisfied identically. Therefore, an extremal cannot be singular for noncircular orbits.

Remark: Defining l_0 by (41) is a mathematical contrivance for the sole purpose of making the constant value of $|Q(\theta)|$ agree with a previous study,[1] in order to use arguments already developed in that work to establish that there are no singular solutions. The actual numerical value of $m(\theta_f)$ and, consequently, l_0 is not known a priori. We could set l_0 equal to a known positive number—e.g., $l_0 = b$—and obtain $|Q(\theta)| = m(\theta_f)(\theta \in \Theta)$. We would then have to modify the arguments of the earlier study to establish the second result on singular extremals.

As we have shown, for l_0 defined by Eq. (41), an extremal is singular if and only if Eq. (42) is valid identically. For circular orbits, $e = 0$, and Eq. (20) shows that an extremal is singular if and only if

$$|q(\theta)| = 1 \tag{43}$$

for each $\theta \in \Theta$ where q is obtained from Eqs. (33). It is known[2,3] that Eq. (43) is valid on a set of positive measure if and only if the constants in Eq. (33) take one of the following sets of values:

$$C_1 = \rho = \alpha = \beta = 0 \quad C_2 = \pm 1 \tag{44}$$

$$C_1 = C_2 = \alpha = 0 \quad \rho = \tfrac{1}{2} \quad \beta = \pm\sqrt{3}/2 \tag{45}$$

Substituting these into Eqs. (33), we see that Lawden's primer must be defined by one of the following expressions:

$$q_1(\theta) = \pm 1, \qquad q_2(\theta) = q_3(\theta) = 0 \tag{46}$$

$$q_1(\theta) = \sin\theta, \qquad q_2(\theta) = \tfrac{1}{2}\cos\theta, \qquad q_3(\theta) = \pm\sqrt{3}/2\cos\theta \tag{47}$$

which are valid for each $\theta \in \Theta$. We thus establish that Eq. (43) holds identically if and only if either Eq. (46) or Eq. (47) do. This provides the third result on singular solutions.

Theorem: An extremal is singular if and only if $e = 0$ and Lawden's primer is given by Eqs. (46) or (47).

This result is valid whether the constant k is positive or zero and is, therefore, independent of whether or not the mass of the spacecraft is assumed constant, and generalizes previous work.[3] In accord with the terminology of that study, we refer to singular extremals defined by Eq. (46) as type-one singular extremals and those defined by Eq. (47) as type-two singular extremals. Each of these two types includes two cases, i. e., two distinct primer functions for which an extremal is singular. We thus have four singular cases. These cases are almost identical to those found for linearized transfer problems using orbital elements in Marec's work (Ref. 14, Chap. 7). He shows that singular solutions cannot occur for $e > 0$ and, although his notation is different, he obtains the equivalent of Eqs. (46) and (47).

If Eq. (46) or (47) is valid on the interval Θ, then any arbitrary measurable function f that satisfies Eq. (14) a.e. on Θ defines a singular extremal control function u through Eq. (16), and via Eq. (6) a singular extremal trajectory. The equation (35) that must be satisfied by the adjoint variable l simplifies because of Eq. (43), but we can directly evaluate this variable from Eqs. (29) and (30) using Eqs. (43), (41), and the fact that $Q = q$ and s is identically zero.

Assuming variable mass ($k > 0$), we obtain

$$l(\theta) = \frac{b}{k}\left[\frac{1}{m(\theta_f)} - \frac{1}{m(\theta)}\right] \tag{48}$$

where Eqs. (5) and (16), and the fact that $e = 0$ determine the mass by

$$m(\theta) = -k\int_{\theta_0}^{\theta} f(\zeta)\, d\zeta + m_0 \tag{49}$$

We have established that singular extremals do exist if the end conditions on the primer q are such that Eq. (46) or (47) holds and we have classified the singular extremals according to four cases and two types. We further classify the singular extremals as follows. If an extremal is singular and the function f that appears in the control function (16) is defined by $f(\theta) = 0$ a.e. on Θ or $f(\theta) = 1$ a.e. on Θ, then we say that the extremal is *marginally singular*. It has been shown that for constant mass, an optimal marginally singular control function is also an optimal nonsingular control function.[3] It can readily be seen from Eqs. (16), (46), and (47) that there are exactly five marginally singular control functions. The trivial marginally singular control function defined by $f(\theta) = 0$ a.e. on Θ is the only singular control in which the four cases defined by Eqs. (46) and (47) are not distinct. We say that a singular extremal is *strictly singular* if it is not marginally singular. A strictly singular extremal is called an *intermediate-thrust* extremal if there is a subset S of Θ of positive measure such that $0 < f(\theta) < 1$ for each $\theta \in S$. It should be noted that there are many strictly singular extremals that are not intermediate-thrust extremals. The author was not aware of a distinction in the literature between these two kinds of singular extremal controls until recently,[3] but the expression "intermediate-thrust arcs" or similar terminology has been used extensively almost as a synonym for "singular regimes" in problems in which the nonsingular extremals are of the form of full thrust and coast. One possible reason is that many researchers did not assume bounded thrust, but instead unbounded thrust and infinite instantaneous impulses to model these types of problems.

Lasalle[15] has shown that for time-optimal control of linear systems, the singular solutions can degenerate in the sense that infinitely many optimal singular trajectories can satisfy the same boundary conditions. This phenomenon has also been observed for linearized fuel-optimal transfer problems in Marec's work (Ref. 14, Chap. 7). It has been proved[3] for the case in which the mass is constant that the intermediate-thrust solutions degenerate in this sense, but a marginally singular solution cannot be included in this type of degeneracy. It is an open question whether or not strictly singular solutions that are not intermediate-thrust solutions can be included in degeneracy. The primary theorem on degeneracy (Ref. 3, Theorem 2.3) is valid also for our problem in which the mass is not assumed constant, but the proof requires considerable modification. In the present context, the theorem, or fourth result, is as follows.

Theorem: If an admissible control function that minimizes Eq. (8), subject to Eqs. (6), (7), (9), and (10) is an intermediate-thrust control function, then this minimization problem has infinitely many intermediate thrust solutions.

We outline the modifications of the original proof as follows. Because of the one-to-one correspondence of Eq. (34), we may replace t in the proof of Theorem 2.3 by θ. Although the state equations (6) are not linear, Eqs. (6a) and (6b) when regarded as equations in y and v, are linear for an arbitrary function m. Using Eqs. (6a) and (6b) in y and v to replace the state equations in x and v in Ref. 3, and replacing u and f, respectively, by u/m and f/m, we define the function f exactly as before, and the proof is then almost exactly the same. It is not difficult to see from Eq. (8) that minimizing $J\,[u]$ is equivalent to minimizing $J\,[u/m]$ become

$$kJ\,[u/m] = -\int_{\theta_0}^{\theta_f} m'(\theta)/m(\theta)\;\mathrm{d}(\theta) = -\log m(\theta_f)$$

$$+\log m_0 = -\log(m_0 - kJ\,[u]) + \log m_0$$

All of the other theorems in the previous work[3] extend to the present situation in which the mass is not considered constant, and the proofs require even less modification than just offered. The most important of these is Theorem 2.6, which can be combined with earlier remarks and stated in the terminology of this present paper thereby establishing our fifth result on singular extremals.

Theorem: An admissible control function that minimizes Eq. (8) subject to Eqs. (6), (7), (9), and (10) is both singular and nonsingular if and only if it is one of the five marginally singular control functions.

Perhaps the most important aspect of this theorem is the following corollary, our sixth result, also found in the previous paper.

Corollary: The problem of finding an admissible control function that minimizes Eq. (8) subject to Eqs. (6), (7), (9), and (10) cannot admit both strictly singular and nonsingular solutions.

We can conclude from these results that there are initial and terminal conditions (9) and (10) such that it is necessary for extremal solutions to be intermediate-thrust solutions and degenerate. There are also conditions where they can be strictly singular and not intermediate-thrust solutions so as to contain arbitrarily many intervals of full thrust and coast. In this situation, it is an unsolved problem whether or not they are degenerate. In theory, these possibilities pose troublesome problems for the construction of fuel-optimal trajectories in a few rare situations where solutions are singular. This is one of the shortcomings of using equations that are based on linearization about a point in circular orbit, although no such problem occurs for other Keplerian orbits. In actual orbital problems, these troublesome situations probably will not occur since Eq. (6b) has been linearized with respect to y, and a number of nonlinear studies that are cited in Ref. 3 do not demonstrate the kinds of singular solutions obtained here. It is likely that regions of degeneracy of singular solutions in this approximate model indicate regions of "flatness of optima" in a more accurate model. For the more accurate nonlinear models, however, the difficulty discovered by Robins[11] in which an optimal trajectory can consist of both singular and nonsingular subarcs may be encountered in some situations.

Structure of the Fuel-Optimal Controls

With the completion of the investigation of the singular extremals, we are able to refine our statements about the variable $\xi(\theta)/b$ of Eqs. (29) and (30) that define the radius of an expanding sphere. We may now assert that the interval Θ is decomposed into finitely many subintervals over which the function ξ is either strictly increasing or constant. It must increase on the intervals where $s(\theta) < 0$. It is a constant on the entire interval Θ if and only if $s(\theta) \geq 0$ for each $\theta \in \Theta$. Switches occur if and only if the expression $|Q(\theta)| - \xi(\theta)/b$ changes sign. Viewed geometrically, the switches occur where the primer $Q(\theta)$ crosses the sphere centered at the origin whose radius is $\xi(\theta)/b$.

The foregoing statements give some geometric insight into the structure of extremal controls. We may also state results concisely in the form of theorems. They fall naturally into two cases.

Theorem: If the orbit is noncircular ($e > 0$), it is necessary that any fuel-optimal control u satisfy the following: Either $Q(\theta) = 0$ for each $\theta \in \Theta$ and $u(\theta) = 0$ a.e. on Θ, or else Q is nonzero except at finitely many values and

$$u(\theta) = -(Q(\theta)/|Q(\theta)|)f(\theta),$$

a.e. on Θ where

$$f(\theta) = \begin{cases} 0, & s(\theta) > 0 \\ 1, & s(\theta) < 0 \end{cases}$$

except at the zeros of s that are finitely many.

Theorem: If the orbit is circular ($e = 0$), it is necessary that any fuel-optimal control u satisfy the following: Either $q(\theta) = 0$ for each $\theta \in \Theta$ and $u(\theta) = 0$ a.e. on Θ, or else q is nonzero except at finitely many values and

$$u(\theta) = -[q(\theta)/|q(\theta)|]f(\theta)$$

a.e. on Θ where

$$f(\theta) = \begin{cases} 0, & s(\theta) > 0 \\ 1, & s(\theta) < 0 \end{cases}$$

except at the zeros of s that are finitely many, or else $s(\theta)$ is identically zero and q is defined by one of the following four formulas:

$$q_1(\theta) = \pm 1, \qquad q_2(\theta) = q_3(\theta) = 0$$

$$q_1(\theta) = \sin\theta, \qquad q_2(\theta) = \tfrac{1}{2}\cos\theta, \quad q_3(\theta) = \pm\sqrt{3}/2\cos\theta$$

and in these cases f is any measurable function satisfying $0 \leq f(\theta) \leq 1$.

In these theorems, Q is given by Eqs. (24–26), q by Eq. (33), and s by Eq. (22) where $Q = q$ for $e = 0$, and l_0 is defined by Eq. (41) for normal problems. For abnormal problems, $l_0 = 0$ and singular solutions cannot occur.

IV. Conclusions

A fixed-duration, bounded-thrust, fuel-optimal rendezvous problem of a spacecraft near an object in arbitrary Keplerian orbit has been investigated using equations of motion in which the Newtonian gravitational force has been linearized about the center of the object. In this formulation the loss of mass of the spacecraft due to propellant consumption, assuming constant exhaust velocity, has not been neglected so that the total equations of state are nonlinear. This approach may be of value in the investigation of problems in which the optimal burn time is too long for the well-established methods of approximating by unbounded thrust with instantaneous impulses to be applicable, or it may be used as a basis for evaluating the effectiveness of these methods.

Because the state equations are nonlinear, we only obtain necessary conditions for fuel-optimal maneuvers. We present two expressions involving the switching function that are useful in determining the structure of optimal solutions. It is found that many known results for the linear problem in which the mass of the spacecraft is regarded constant also generalize to this problem. Among these are some recently published results about singular trajectories. These trajectories, which only occur for certain circular orbits, exhibit several interesting properties.

Nonsingular optimal trajectories must be of the form of intervals of full thrust and coast separated by a finite number of switches. If the spacecraft mass is assumed constant, these switches are determined from points of intersection of a primer vector locus with a unit sphere; otherwise, this geometric picture changes and the switches are determined from the intersections of this primer locus with an expanding sphere of monotone nondecreasing radius.

This work applies to orbits that are circular, elliptical, parabolic, or hyperbolic. Possible applications of this kind of analysis include rendezvous near a satellite, space station, or small asteroid or comet. A second type of application includes rendezvous with a fictitious position and velocity in order to correct for accumulated errors from a specified point in orbit. This second type of application could involve very general space missions in which segments of the total trajectory are Keplerian.

References

[1]Carter, T. and Humi, M., "Fuel-Optimal Rendezvous near a Point in General Keplerian Orbit," *Journal of Guidance, Control, and Dynamics*, Vol. 10, Nov.–Dec. 1987, pp. 567–573.

[2]Carter, T., "Fuel-Optimal Maneuvers of a Spacecraft Relative to a Point in Circular Orbit," *Journal of Guidance, Control, and Dynamics*, Vol. 7, 1984, pp. 710–716.

[3]Carter, T., "Singular Fuel-Optimal Space Trajectories Based on a Linearization about a Point in Circular Orbit," *Journal of Optimization Theory and Applications*, Vol. 54, Sept. 1987, pp. 447–470.

[4]De Vries, J. P., "Elliptic Elements in Terms of Small Increments of Position and Velocity Components," *AIAA Journal*, Vol. 1, 1963, pp. 2626–2629.

[5]Tschauner, J. and Hempel, P., "Rendezvous zu einem in elliptischer Bahn umlaufenden Ziel," *Astronautica Acta*, Vol. 11, March 1965, pp. 104–109.

[6]Shulman, Y. and Scott, J., "Terminal Rendezvous for Elliptical Orbits," AIAA Paper 66-533, June 1966.

[7]Euler, E. A., "Optimal Low-Thrust Rendezvous Control," *AIAA Journal*, Vol. 7, 1969, pp. 1140–1144.

[8]Lawden, D. F., "Fundamentals of Space Navigation," *Journal of the British Interplanetary Society*, Vol. 13, March 1954, pp. 87–101.

[9]Tschauner, J., "Elliptic Orbit Rendezvous," *AIAA Journal*, Vol. 5, 1967, pp. 1110–1113.

[10]Lee, E. B. and Markus, L., *Foundations of Optimal Control Theory*, Wiley, New York, 1967.

[11]Robbins, H. M., "Optimal Rocket Trajectories with Subarcs of Intermediate Thrust," 17th Congress of the International Astronautical Federation, Madrid, Oct. 1966.

[12]Lawden, D. F., *Optimal Trajectories for Space Navigation*, Butterworths, London, 1963.

[13]Carter, T., "How Many Intersections Can a Helical Curve Have with the Unit Sphere During One Period," *The American Mathematical Monthly*, Vol. 93, Jan. 1986, pp. 41–44.

[14]Marec, J. P., *Optimal Space Trajectories*, Elsevier, New York, 1979.

[15]LaSalle, J. P., *The Time Optimal Control Problem: Contributions to the Theory of Nonlinear Oscillations*, Vol. 5, Princeton Univ. Press, Princeton, NJ, 1960, pp. 1–24.