Globally Optimal Impulsive Transfers via Green's Theorem

G.A. Hazelrigg Jr.*
National Science Foundation, Washington, D.C.

For certain classes of trajectories the cost function (characteristic velocity) can be written as a "quasilinear" function of the change in state. In the case presented, impulsive transfers between coplanar, coaxial orbits with transfer time and angle unrestricted, Green's theorem can be used to determine the optimal transfer between given terminal states. This is done in a manner which places no restrictions on the number of impulses used and leads to globally optimal results. These results are used to show that the Hohmann transfer and the biparabolic transfer provide global minima in their respective regions. The regions in which monoelliptic and biparabolic trajectories are globally optimal are also defined for elliptic terminal states. The results are applicable to the case in which restrictions are placed on the radius of closest approach or greatest recession from the center of the force field.

Nomenclature

A,B,C,D,E,F	= constants of integration
a	= semimajor axis (also defined for a hyperbola)
e	= orbital eccentricity
f	= true anomaly
H	= Hamiltonian (constant of motion)
K	= switch function [defined by Eq. (5)]
ℓ	= semilatus rectum
m	= spacecraft mass
$\bar{p}$	= primer vector (adjoint variables)
$\bar{r}$	= radius vector, r = magnitude
r_a	= apoapse radius (also defined for a hyperbola)
r_p	= periapse radius
V	= velocity
V_j	= effective jet velocity
ΔV	= characteristic velocity
X	= inverse of apoapse radius
β	= propellant mass flow rate
θ	= longitude
λ,μ,ν	= components of $\bar{p}$
μ	= variable defined by Eq. (39)
ν	= variable defined by Eq. (40)
ξ	= variable defined by Eq. (35)
σ	= adjoint variable corresponding to mass
$\phi,\psi,$	= cost functions defined by Eqs. (26) and (27)
Ω	= latitude
ω	= fundamental function
$\tilde{\omega}$	= longitude of periapse

Introduction

ONE of the earliest studies of orbital transfers was the classical work of Hohmann[1] in which the author proposed a bitangent elliptic transfer of angle π between two coplanar, circular orbits in an inverse square force field. The transfer which now bears his name had long been assumed to be the trajectory which provides the absolute minimum impulse requirement between terminal orbits (globally optimal). In 1959, however, Shternfeld,[2] Hoelker and Silber[3] and Edelbaum[4] showed that when the terminal orbits are sufficiently far apart three-impulse, bielliptic trajectories can provide transfers of lower cost than the two-impulse Hohmann. Later, Ting[5] proved that when the orbits are sufficiently close, the Hohmann trajectory is locally optimal (optimal with respect to near-Hohmann transfers). He then showed[6] that for time-open, angle-open transfers all trajectories of four or greater impulses could be reduced to two- or three-impulse transfers with reduced cost. The reduction to two impulses occurred when the orbits were sufficiently close and it was noted that a proof of the global optimality of the Hohmann transfer within the restricted, two-impulse class of transfers would now suffice to prove the global optimality of the Hohmann transfer without restriction on the number of impulses. In 1963, Barrar[7] supplied the necessary proof, thus concluding that the Hohmann transfer was the globally optimal impulsive transfer. In 1964, McIntyre[8] showed that any finite thrust trajectory could be approximated to any desired degree of closeness by an impulsive trajectory of less or equal cost. Thus, the restriction on thrust level was removed and it became clear that the Hohmann transfer was without restriction the globally optimal transfer between near orbits.

In 1965, based on the preliminary efforts of Breakwell[9] and Contensou,[10] Moyer[11] analyzed the circle-to-ellipse transfer using orbital elements. This formulation enabled him to consider transfers using n impulses and resulted in providing the global optimality of both the Hohmann-type transfer and the biparabolic transfer for circle-to-ellipse terminal orbits.

Still later, in 1979, Marec[12] demonstrated the optimality of the Hohmann transfer using a hodograph analysis.

The work which follows, completed in 1968,[13] offers still another proof of the global optimality of the Hohmann transfer (and biparabolic) which, the author feels, is more general than the above approaches and leads to the result in a more direct manner. Also included are globally optimal transfers from elliptic (or circular) orbits to hyperbolic orbits which might require up to four impulses. Furthermore, the method presented applies when restrictions are placed on the radius of closest approach or greatest recession from the center of the force field.

To prove that Hohmann-type or biparabolic transfers between elliptic (or circular) orbits are globally optimal without restriction, one must first recognize the following:

1) Any globally optimal time-open, angle-open transfer must be equal to or better than any other trajectory for which either time or angle or both are constrained.

Submitted June 10, 1982; revision received June 22, 1983.

*Policy Analyst, Division of Policy Research and Analysis. Member AIAA.

2) Any globally optimal trajectory must satisfy all necessary conditions required of locally optimal trajectories, i.e., they must be locally stationary or lie on a boundary.

The method used in the following work makes use of Green's theorem in a coordinate system in which the cost function can be expressed linearly. Similar work has been published by Lee and Markus[14] in which they made use of a slightly different coordinate system and failed to obtain all of the global results presented.

Equations of Motion and Necessary Conditions

Normalizing distance and time so that the gravitational constant is unity, the equations of motion of a space vehicle are

$$\ddot{\bar{r}} = -\frac{\bar{r}}{r^3} + \frac{V_j \beta}{m}\frac{\bar{p}}{p} \tag{1}$$

where the direction of $\bar{p}$ corresponds to the direction of the thrust. Lawden[15] has shown that $\bar{p}$ corresponds to the adjoint variables associated with the velocity and calls $\bar{p}$ the primer vector. The adjoint variables obey the equation

$$\ddot{\bar{p}} = \bar{p}\cdot\nabla\left(-\frac{\bar{r}}{r^3}\right) \tag{2}$$

and the mass varies according to

$$\dot{m} = -\beta \tag{3}$$

Minimizing propellant consumption via the Pontryagin maximum principle requires maximizing the Hamiltonian with respect to the control. For this system the Hamiltonian can be written

$$H = \beta K - \frac{\bar{p}\cdot\bar{r}}{r^3} - \dot{\bar{p}}\cdot\dot{\bar{r}} \tag{4}$$

where

$$K = \frac{V_j p}{m} - \sigma \tag{5}$$

is called the switch function. σ is the adjoint variable corresponding to the mass and obeys the equation

$$\dot{\sigma} = \frac{V_j \beta}{m^2} p \tag{6}$$

Maximizing H with respect to β gives $\beta = \beta_{max}$ when $K \geq 0$, and $\beta = 0$ when $K \leq 0$. If β is unbounded, $\beta_{max} = \infty$, then when $K \geq 0$, $\beta = \infty$. The Hamiltonian represents the first integral for time-invariant systems, hence, H is a constant. Since the terms $\dot{\bar{p}}\cdot\dot{\bar{r}}$ and $\bar{p}\cdot\bar{r}/r^3$ must be finite for finite nonzero radius, velocity, and p, the Hamiltonian can remain constant and the previous condition that for $K \geq 0$, $\beta = \infty$ can be satisfied only if K is never greater than zero. It follows that for an optimal trajectory $K \leq 0$ at all times and the periods of propulsion occur when $K = 0$. There are two possibilities: $K = 0$ at a maximum point or $K \equiv 0$ over some finite interval. If the former case holds, in order to obtain any propulsion at all $\beta = \infty$. For the latter case, called a singular arc, β is undetermined. Points where $K = 0$ arise when, for time-open trajectories,

$$\bar{p}\cdot\dot{\bar{p}} = 0, \qquad \bar{p}\cdot\ddot{\bar{p}} < 0 \tag{7}$$

i.e., p is at a maximum point. This is called an impulse and is a point at which the control is such that the velocity of the vehicle may change in a discontinuous manner but $\bar{r}$ remains continuous. The necessary condition (7) and the following necessary conditions at an impulse are derivable from the constants of the motion[16]:

1) $\bar{p}$ and $\dot{\bar{p}}$ are continuous across an impulse.
2) The impulse is in the direction of $\bar{p}$.
3) H is continuous across an impulse.
4) The product $m\sigma$ is constant.

Edelbaum[17] points out that the only singular subarc for the coplanar time-open problem has been discussed by Lawden,[15,18,19] and shown to be nonoptimum by several authors.[20-24] Thus, if β is unconstrained, its optimum value is infinity. Transfers which occur with $\beta = 0$ or $\beta = \infty$ are called impulsive transfers. This argument suffices to prove that for time-open transfers any finite-thrust trajectory can be equaled or bettered by an impulsive trajectory.

Solution of the Primer on Coast Arcs

The solution for the primer on a circular ($e = 0$) coast arc is given by Lawden[15] as follows:

$$\bar{p} = \begin{bmatrix} \lambda \\ \mu \\ \nu \end{bmatrix} = \begin{bmatrix} A\cos f + B\sin f + 2C \\ 2B\cos f - 2A\sin f - 3Cf + D \\ E\cos f + F\sin f \end{bmatrix} \tag{8}$$

where A, B, C, D, E, and F are constants and

$$f = \theta - \tilde{\omega} \tag{9}$$

is the true anomaly.

In the case where $e \neq 0$, the solution becomes

$$\bar{p} = \begin{bmatrix} \lambda \\ \mu \\ \nu \end{bmatrix} = \begin{bmatrix} A\cos f + Be\sin f + CI_1 \\ -A\sin f + B(1 + e\cos f) + \dfrac{D - A\sin f}{1 + e\cos f} + CI_2 \\ (1 + e\cos f)^{-1}(E\cos f + F\sin f) \end{bmatrix} \tag{10}$$

where

$$I_1 = \sin f \int \frac{df}{\sin^2 f(1 + e\cos f)^2} \tag{11}$$

$$I_2 = \frac{\cot f}{e(1 + e\cos f)} + \frac{1 + e\cos f}{e\sin f} I_1 \tag{12}$$

The derivative of the primer vector is, for $e = 0$,

$$\dot{\bar{p}} = \begin{bmatrix} \dot{\lambda} \\ \dot{\mu} \\ \dot{\nu} \end{bmatrix} = \begin{bmatrix} a^{-3/2}(A\sin f - B\cos f + 3Cf - D) \\ -a^{-3/2}(A\cos f + B\sin f + C) \\ a^{-3/2}(F\cos f - E\sin f) \end{bmatrix} \tag{13}$$

and, for $e \neq 0$,

$$\dot{\bar{p}} = \begin{bmatrix} \dot{\lambda} \\ \dot{\mu} \\ \dot{\nu} \end{bmatrix} = \begin{bmatrix} \dfrac{\sqrt{\ell}}{r^2}\left(\dfrac{A\sin f - D}{1 + e\cos f} - B + CI_3\right) \\ \ell^{-3/2}\{-A(e + \cos f) + De\sin f + C\cos f\} \\ \ell^{-3/2}\{(e + \cos f)F - E\sin f\} \end{bmatrix} \tag{14}$$

where

$$I_3 = \frac{e\sin f - \cos f}{e\sin f(1+e\cos f)^2} - \frac{I_1}{e}\csc f \tag{15}$$

The constant C is associated with the Hamiltonian or first integral of the motion and for time-open trajectories $C=0$. Furthermore, reducing the problem to one in two dimensions requires $E=F=0$ since ν and $\dot{\nu}$ are zero. Lastly, the specification of angle-open leads to the fact $A=0$.[25] $\bar{p}$ is thus reduced to the two-dimensional vector for $e=0$.

$$\bar{p} = \begin{bmatrix} \lambda \\ \mu \end{bmatrix} = \begin{bmatrix} B\sin f \\ 2B\cos f + D \end{bmatrix} \tag{16}$$

and, for $e \neq 0$,

$$\bar{p} = \begin{bmatrix} \lambda \\ \mu \end{bmatrix} = \begin{bmatrix} Be\sin f \\ B(1+e\cos f) + \dfrac{D}{1+e\cos f} \end{bmatrix} \tag{17}$$

$\dot{\bar{p}}$ is, for $e=0$,

$$\dot{\bar{p}} = \begin{bmatrix} \dot{\lambda} \\ \dot{\mu} \end{bmatrix} = \begin{bmatrix} a^{-3/2}(-B\cos f - D) \\ -a^{-3/2}B\sin f \end{bmatrix} \tag{18}$$

and, for $e \neq 0$,

$$\dot{\bar{p}} = \begin{bmatrix} \dot{\lambda} \\ \dot{\mu} \end{bmatrix} = \begin{bmatrix} \sqrt{\ell}/r^2\left(-\dfrac{D}{1+e\cos f} - B\right) \\ \ell^{-3/2}De\sin f \end{bmatrix} \tag{19}$$

Condition (7) can be satisfied only for $f=0, \pi$ when $e=0$. It is readily concluded that a circular orbit must be entered or exited at an apse point on the adjacent conic. Furthermore, any ellipse entered at an apse point must be exited at an apse point and vice versa. Clearly, if a terminal or intermediate orbit is circular, a necessary condition is that all other orbits, with the possible exception of a terminal parabola or hyperbola, be coaxial to the point of entrance or exit on the circular orbit. Appendix A shows that for open-time open-angle transfers all orbits entered or exited at finite radius, if optimally oriented will be coaxial. However, it is possible to change the orientation of a parabola or hyperbola by an impulse at infinite radius for zero cost.

The remainder of this article will deal with transfers between adjacent orbits by means of impulses at the apse points only except as allowance is made for impulses at infinite radius in the case of hyperbolic and parabolic trajectories. It is, therefore, assumed that all orbits are coaxial except terminal hyperbolic arcs which may assume some other orientation. In that case, the orientation is left free to be optimized.

A final point must be clarified. In contradiction to the work of Lawden,[15] an impulse is possible at infinite radius on a parabolic or hyperbolic arc. As the radius tends to infinity, p tends to some fixed value. Hence, in the limit $\dot{p}$ tends to zero and if $p=1$, condition (7) is automatically fulfilled. With $D=0$ [Eq. (8)] p may become unity at $r=\infty$ since $(1+e\cos f)=0$ also.

It is assumed than on rectilinear orbits, impulses occur only at $r=0$ (periapse) or $r=r_{\max}$ (apoapse). Impulses at infinite radius are of infinitesimal size. Therefore, they change angular momentum (r_p) but not the energy $(r_a + r_p = \text{const})$.

Mapping the Accessible States

The above necessary conditions for optimal time-open, angle-open transfers between optimally oriented orbits have eliminated all but coaxial orbits as intermediate and terminal arcs except for the possibility of changing orientation via an infinitesimal impulse at infinity. Hence, only two variables are necessary to describe the orbit of the vehicle, e.g., eccentricity and semimajor axis or energy and angular momentum. These states may be mapped on the desired plane and the allowable, optimal changes in state are dictated by the primer vector. It is desirable to choose a set of axes such that the direction in which a change of state is made is independent of the magnitude of the impulse producing the change. Such a set of axes is $1/r_a$ vs r_p, shown in Fig. 1. (For a hyperbola $r_a<0$ and $|r_a|$ is the distance from the periapse to the vacant focus.) The necessary conditions and the primer vector indicate that in the elliptic region only vertical or horizontal transitions are allowable on this plane. In the hyperbolic region, vertical and constant energy transfers only are allowed. That is, impulses must occur at either the apoapse (or infinity) or the periapse and must be directed perpendicular to the line of apsides (except at infinity). Thus, on ellipses impulses at the apoapse change only the periapse and vice versa. Furthermore, on this plot there are two unique maps of cost (ΔV), one for changes in X and one for changes in r_p. These maps are independent of the initial or final orbit states and are a direct measure of the cost of changing state.

Consider now only the elliptic region. To this point almost any path comprised only of vertical and horizontal segments and connecting an initial and final state on the X-r_p plane satisfies some of the necessary conditions of the calculus of variations (Fig. 2). Further restrictions on allowable transformations between states 1 and 2 may be made by invoking the required continuity of $\dot{p}$. However, an eloquent treatment of the linear problem has been presented by Miele[26] and, although this case is not linear in the strict sense, results can nonetheless be obtained.

Curves of cost (ΔV) can be constructed easily for changes in X and r_p as follows: At any point on an orbit

$$V = \sqrt{\frac{2}{r} - \frac{2}{a}} \tag{20}$$

Thus, changes in X which occur at periapse require

$$\Delta V_x = \sqrt{2}\left\{\sqrt{\frac{1}{r_p(1+X^0 r_p)}} - \sqrt{\frac{1}{r_p(1+X^1 r_p)}}\right\} \tag{21}$$

taking the axis $X=0$ as a reference line, $\Delta V_x(X=0)=0$

$$\Delta V_x = \sqrt{\frac{2}{r_p}}\left\{1 - \sqrt{\frac{1}{1+Xr_p}}\right\} \tag{22}$$

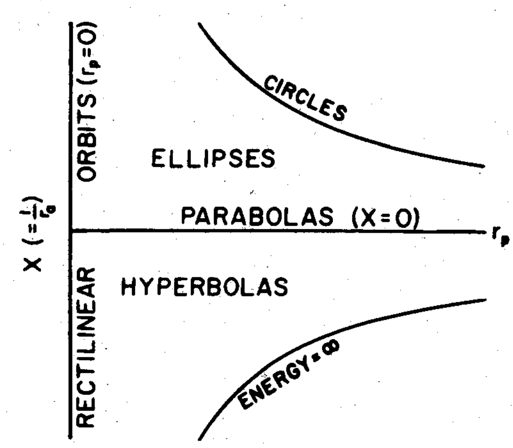

Fig. 1 X-r_p **state plane.**

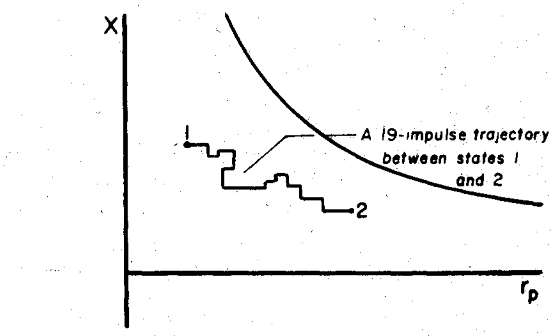

Fig. 2 Example of an impulsive transfer satisfying optimality conditions imposed by $\bar{p}$ on the state plane.

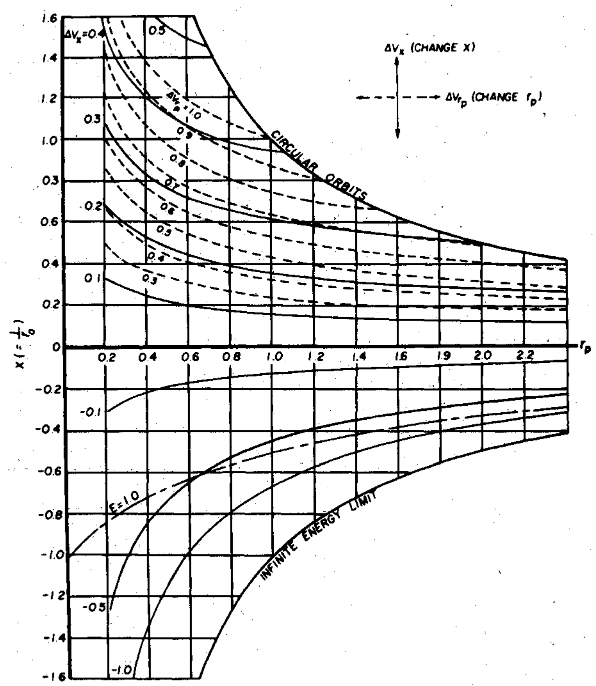

Fig. 3 Map of cost (ΔV) for transfers on the X-r_p plane.

For changes in r_p which occur at the apoapse,

$$\Delta V_{r_p} = \sqrt{2X}\left\{\sqrt{\frac{Xr_p^I}{1+Xr_p^I}} - \sqrt{\frac{Xr_p^0}{1+Xr_p^0}}\right\} \tag{23}$$

and taking the axis $r_p=0$ as a reference line, $\Delta V_{r_p}(r_p=0)=0$

$$\Delta V_{r_p} = X\sqrt{\frac{2r_p}{1+Xr_p}} \tag{24}$$

except for $X \leq 0$ when $\Delta V_{r_p}=0$ for any change in r_p (occurs at $r=\infty$). These map as shown in Fig. 3.

The total cost required to transfer from state I to state F along a path Q is

$$\Delta V = \int_{IQF} (\phi \mathrm{d}r_p + \psi \mathrm{d}X) \tag{25}$$

where

$$\phi = \frac{\partial \Delta V_{rp}}{\partial r_p} \tag{26}$$

$$\psi = \frac{\partial \Delta V_x}{\partial X} \tag{27}$$

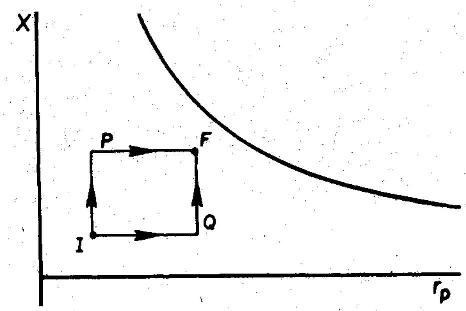

Fig. 4 Comparison of two paths from I to F, F above and right of I.

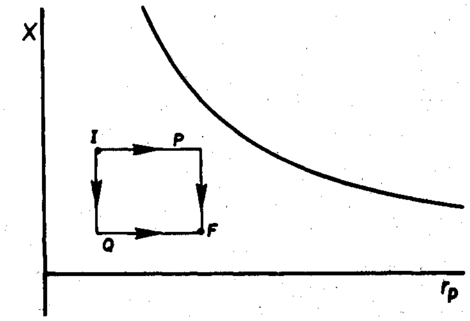

Fig. 5 Comparison of two paths from I to F, F below and right of I.

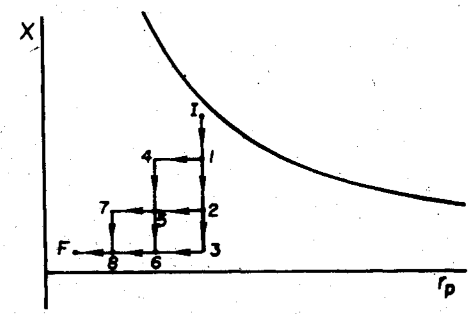

Fig. 6 Comparison of paths I to F, F below and left of I.

This equation holds because only horizontal and vertical paths in a single direction are considered (all other paths are eliminated by optimality conditions). Using Eqs. (22) and (24), (26) and (27) become

$$\phi = \frac{X}{\sqrt{2r_p}}(1+Xr_p)^{-3/2} \tag{28}$$

$$\psi = \sqrt{\frac{r_p}{2}}(1+Xr_p)^{-3/2} \tag{29}$$

Comparing path Q to another path P yields, using Green's theorem,

$$\delta\Delta V = \oint_{IQFPI}(\phi \mathrm{d}r_p + \psi \mathrm{d}X) = \int\int_\alpha \left(\frac{\partial \psi}{\partial r_p} - \frac{\partial \phi}{\partial X}\right)\mathrm{d}r_p \mathrm{d}x \tag{30}$$

$$\delta\Delta V = \int\int_\alpha \omega(r_p, X)\,\mathrm{d}r_p \mathrm{d}X \tag{31}$$

where ω is the fundamental function

$$\omega(r_p, X) = \frac{\partial \psi}{\partial r_p} - \frac{\partial \phi}{\partial X} \tag{32}$$

Since the cost is equal to the magnitude of the change of ΔV, one must take care to choose the proper signs.

$$\frac{\partial \psi}{\partial r_p} = \frac{1}{2\sqrt{2r_p}}\left[\frac{1-2Xr_p}{(1+Xr_p)^{5/2}}\right] \tag{33}$$

$$\frac{\partial \phi}{\partial X} = \frac{1}{2\sqrt{2r_p}}\left[\frac{2-Xr_p}{(1+Xr_p)^{5/2}}\right] \tag{34}$$

The group

$$\xi = \frac{(1+Xr_p)^{-5/2}}{2\sqrt{2r_p}} > 0 \tag{35}$$

Now, if $\omega > 0$ everywhere, $\delta\Delta V > 0$ and path Q is worse than path P. Consider the comparison of paths P and Q, Fig. 4. In this case,

$$\omega = \frac{\partial \psi}{\partial r_p} - \frac{\partial \phi}{\partial X} = -\xi(1+Xr_p) \tag{36}$$

Since $\xi > 0$ and $(1+Xr_p) > 0$, ω is always negative. Thus, path Q is always better than path P. Now consider the comparison of Fig. 5. Since the vertical elements oppose a positive $\mathrm{d}X$, the sign of ψ must change. ω becomes

$$\omega = -\xi(Xr_p - 1) \tag{37}$$

But $(Xr_p - 1) < 0$, hence $\omega > 0$ and path Q is better than path P. If the points I and F are switched, in either case, the sign of ω reverses and Q remains the minimizing path. Clearly, the cost of transferring from I to F equals the cost of transferring from F to I and the optimal path on the X-r_p plane remains the same. The above theorems apply only in the elliptic region of the X-r_p plane. The axes and the $Xr_p = 1$ lines are the boundaries of this region. Transfers in the hyperbolic region occur optimally with impulses at infinite radius and $r_p = 0$, only and at zero cost.

Thus far, the optimal two-impulse transfer between any two points in the elliptic region has been found. It is now possible to derive some theorems on n-impulse trajectories. Consider the possible transfers from I to F shown in Fig. 6. I and F need not be terminal points, but can represent two intermediate states of a transfer. In this case, both $\mathrm{d}r_p$ and $\mathrm{d}X$ are always of the same sign and the result of Eq. (35) applies yielding $I3F$ as the optimal path. If F is moved to the right of I (Fig. 7), the result of Eq. (37) applies and $I3F$ remains the optimal path.

Now consider the case when the sign of $\mathrm{d}r_p$ or $\mathrm{d}X$ is not constant (Fig. 8). Consider the transfer from I to F via $I123467F$. The segment 34 has a negative $\mathrm{d}r_p$ and hence 1234 is not included in the above cases. However, the transfer from 3 to 6 has been treated, the optimal path being 376. Hence, the transfer 37 is clearly more optimal than 3467. This eliminates the loop 3467 and the problem is reduced to one of the above cases. By induction, this theory can be extended to any number and arrangement of impulses.

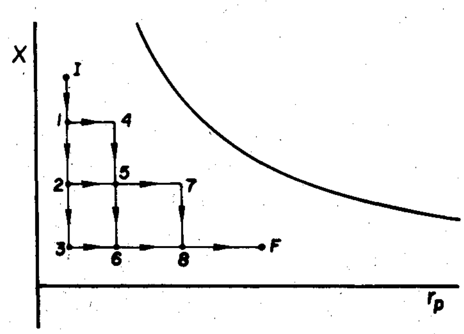

Fig. 7 Comparison of paths I to F, F below and right of I.

From the above, the following theorem can be stated regarding a transfer from I to F (Fig. 9):

1) No state in the shaded region or on its boundaries except for the line $I1F$ can be an intermediate point of an optimal trajectory between states I and F. There is an interesting corollary to this theorem. Let state F lie on the $X=0$ axis. Then theorem 1 becomes: The globally optimal maneuver to parabolic escape is one which follows a vertical path $I1$ from I to $X=0$. A one-impulse transfer from any elliptical orbit to parabolic escape will provide the globally minimal cost. This transfer is not unique, however, as it can be broken into a transfer of any number of impulses with coast phases representable as intermediate points on the line $I1$ at no additional cost (ΔV). It is interesting to note that this infinity of optimal trajectories each of which requires the globally minimal cost maps into a unique line in X-r_p space.

One can also draw additional conclusions regarding the design of an optimal transfer in the unshaded region below the $1F$ line. First, the globally minimal, unconstrained transfer from any elliptical orbit to any hyperbolic orbit on

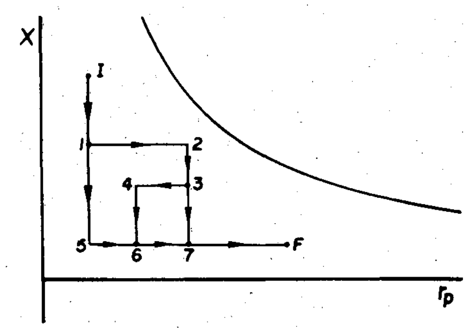

Fig. 8 Comparison of paths I to F with nonmonotone changes in state.

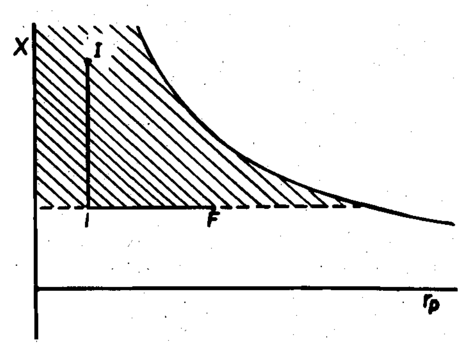

Fig. 9 Nonoptimal region for intermediate states on state plane.

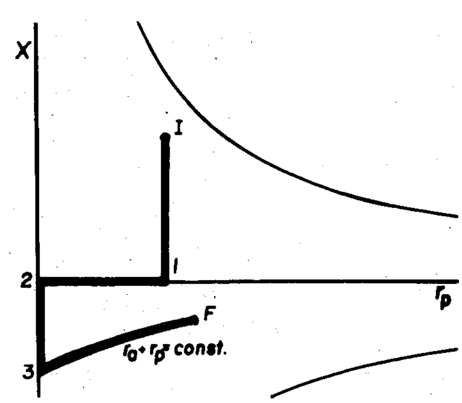

Fig. 10 Optimal transfer to a hyperbolic state.

the X-r_p plane occurs, as shown above, according to Fig. 10. $I1$ is a tangential impulse at perapse, 12 is an impulse at infinity, 23 at $r=0$ and $3F$ at infinity again. Only the transfer $I1$ requires a finite cost. In fact, in the hyperbolic region $\phi=0$ and ψ remains as before. Thus, for transfers in the hyperbolic region ($\mathrm{d}X<0$),

$$\omega=-\xi(1-2Xr_p) \tag{38}$$

Since $X<0$, $\omega<0$ throughout the region and the optimal transfer (changes in X) occurs as close as possible to the $r_p=0$ line.

Now consider the transfer from 1 to F of Fig. 9. In general, only the region above the $1F$ line has been excluded and the transfer below can take on as yet any shape, for example, Fig. 11. Two things become apparent: The inclusion of any intermediate state to the left of 1 or the right of F automatically requires the existence of points a and b on vertical lines from 1 and F. But it was shown above that the best transfer from 1 to a or b to f is along the vertical path $1a$ or bF. Thus: 2) For transfers which occur entirely in the elliptical region, no optimal transfer between I and F can possess an intermediate state which lies outside the region bounded by vertical lines passing through state points I and F. Second, the transfer between points a and b can be constructed of a series of rectangles, e.g., Fig. 12. Consider the transfer qrs. This is nonoptimal and can be bettered by the transfer of Fig. 13, $qr's$. Similary, tuv becomes $tu'v$ (Fig. 14). By induction, therefore, the optimal transfer from a to b must take the form shown in Fig. 15.

There remains only one family of optimal trajectories in which the globally optimal trajectory must exist shown in Fig. 16. A particular member of this family may be characterized by the value of X_a where $0\le X_a\le X_F$. The case when $X_a=X_F$ is called the Hohmann and $X_a=0$ the biparabolic transfer. Normalizing X_a to X_F leads to a parameter which is interpreted more easily:

$$\mu=X_a/X_F \tag{39}$$

where $0\le\mu\le 1$. It is also convenient to define another parameter ν,

$$\nu=r_{PF}X_F \tag{40}$$

which determines the eccentricity of the final orbit.

$$e=\frac{1-\nu}{1+\nu} \tag{41}$$

μ and ν, as used here, should not be confused with the components of the primer vector as presented earlier. Figures 17a and b show the cost (ΔV) of transfering from state 1 to F as a function of μ and r_{PF}, assuming that $r_{PI}=1$. Clearly, for r_{PF} sufficiently small, $\mu=1$ yields the most economical transfer whereas for values of r_{PF} greater than r^*, which satisfies

$$r^{*3/2}-\frac{1+2\nu-\sqrt{1+\nu}}{\sqrt{1+\nu}-1}r^*+\nu r^{*\frac{1}{2}}+\nu=0 \tag{42}$$

$\mu=0$ provides the most economical transfer. The value $r^*(\nu)$ is the exact point at which $\mu=0$ and $\mu=1$ provide transfers of identical cost. Figures 18 and 19 show the crossover from Hohmann to biparabolic transfers for different values of ν.

With some degree of care, the theory developed above and the conclusions derived therefrom apply equally well to transfers with $r_{PF}<r_{PI}$.

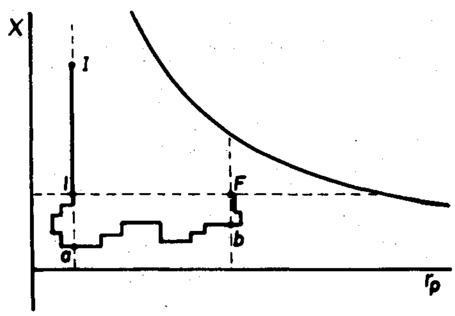

Fig. 11 Example of a transfer from 1 to F.

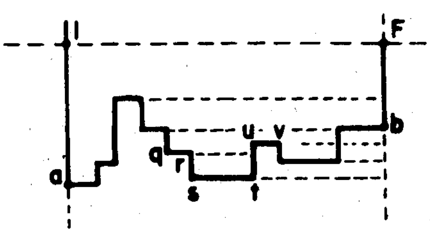

Fig. 12 Breakdown of transfer from 1 to F.

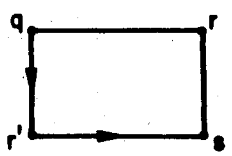

Fig. 13 Improvement of transfer from q to s.

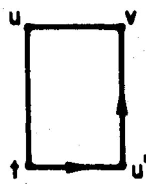

Fig. 14 Improvement of transfer from t to v.

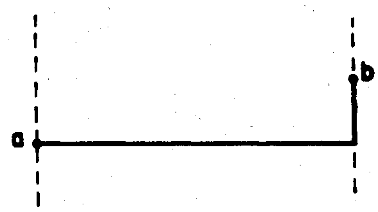

Fig. 15 Optimal transfer from a to b.

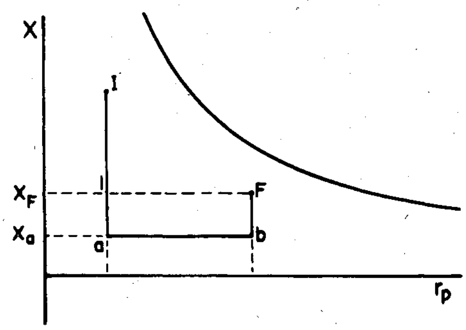

Fig. 16 Remaining family of optimal transfers from I to F.

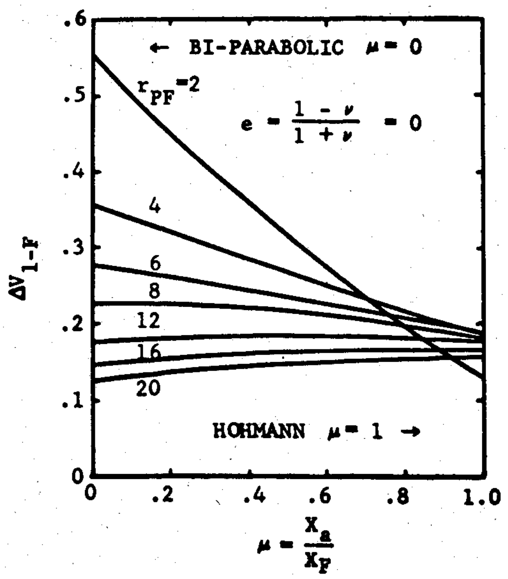

Fig. 17a Cost of transfers as a function of μ.

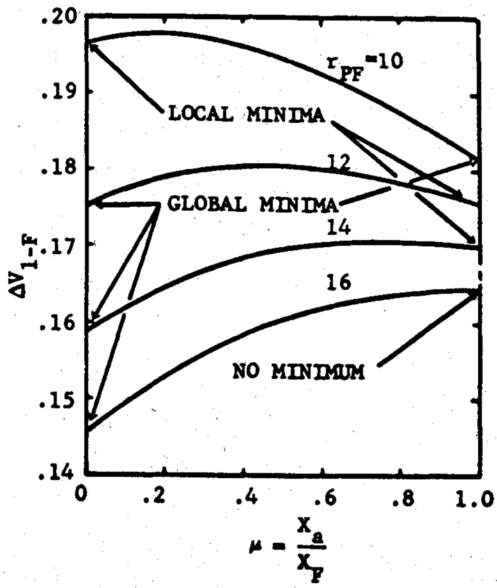

Fig. 17b Cost of transfers as a function of μ—expanded scale.

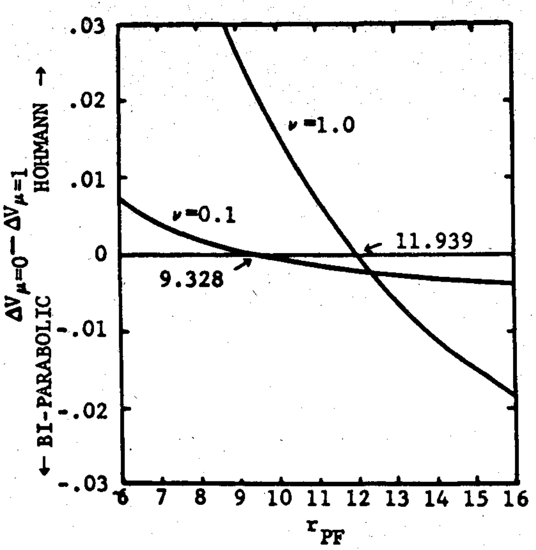

Fig. 18 Differential cost between a Hohmann and biparabolic transfer.

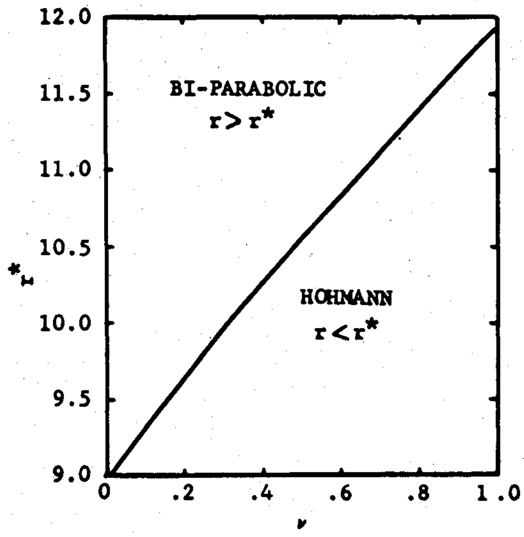

Fig. 19 Crossover between monoelliptic and biparabolic transfers.

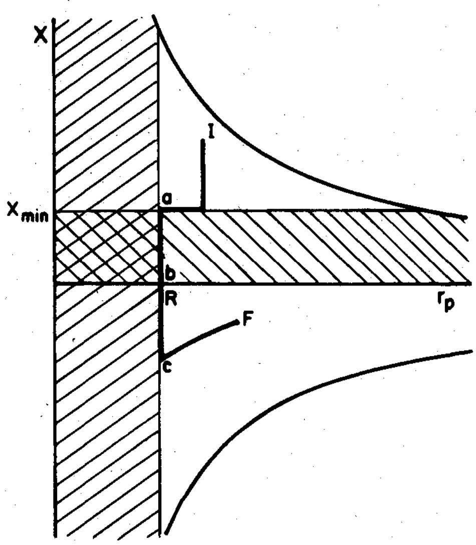

Fig. 20 Constrained optimal transfer to a hyperbolic state.

This concludes the determination of the globally optimal time-open, angle-open transfer between coaxial terminal orbits and a proof of the global optimality of the Hohmann transfer.

Radius Constraints

The work of the previous section applies readily to the case when the radius is constrained to be within a given range at all times. Suppose, for example, that the presence of a planet precludes the possiblity that $r_p < R$, the radius of the planet's atmosphere and because of time limitations $X > X_{\min}$ at all times. The corresponding areas of the X-r_p plane become forbidden zones in which no intermediate state may exist (Fig. 20). The optimum transfer from I to F obeys the rules set down in the previous section but now the transfer from a to c must occur in one impulse whereas without the constraint on X, the region a to b could be broken into an arbitrary number of impulses.

Transfers between elliptic orbits such as that of Fig. 21 have two possible candidates for the globally optimal transfer: paths IaF and $IabcF$. A numerical comparison of the cost for each immediately discloses the better. Curves of the type of Fig. 7 help to gain an exact understanding of the variation of the cost with the transfer path.

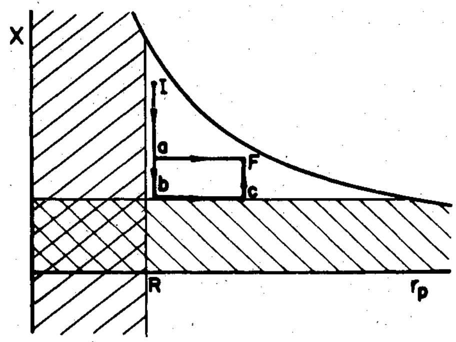

Fig. 21 Constrained optimal transfer to an elliptic state.

Conclusions

Primer vector theory has been applied to the problem of minimum impulse orbital transfer, with emphasis on the class of transfers that is of unconstrained duration and between terminal orbits of unconstrained orientation. The principal result obtained reconfirms the global optimality of the Hohmann-type and biparabolic transfers. The method used, however, is both more concise and less restrictive than previous methods. In particular, it imposes no restrictions on the relative sizes of the terminal orbits, and it applies in the case in which restrictions are imposed on the radius of closest approach or greatest recession from the center of the force field. It is also shown that, for time-open transfers, any finite-thrust transfer can be equaled or bettered by an impulsive transfer.

Appendix A: Derivation of the Transversality Equation for Optimal, Time-Open, Angle-Open Transfers Between Optimally Oriented Orbits

The general form of the transversality equation in spherical coordinates is

$$\lambda_u du + r\cos\phi\lambda_\omega d\omega + r\lambda_\Omega d\Omega + \lambda_r dr + \lambda_\theta d\theta + \lambda_\phi d\phi - Cdt\Big|_{t_0}^{t_f} = 0 \tag{A1}$$

where $u = \dot{r}$ and

$$C = r(\Omega^2 + \omega^2\cos^2\phi)C' \tag{A2}$$

and the λ's are the adjoint variables on the designated state variables. Note that if the latitude is unconstrained, $d\phi \neq 0$ then $\lambda_\phi = 0$. Thus $F = E = 0$ and $\lambda_\Omega = 0$ also. Hence, two-dimensional transfers are conducted at optimal latitude, $\phi = 0$.

Two Dimensions

The transversality equation is

$$\lambda_{uf}du_f + r_f\lambda_{\omega f}d\omega_f + \lambda_{rf}d_{rf} + \lambda_{\theta f}d\theta_f - Cdt_f - \lambda_{u0}du_0 - r_0\lambda_{\omega 0} - \lambda_{r0}dr_0 - \lambda_{\theta 0}d\theta_0 + Cdt_0 = 0 \tag{A3}$$

C corresponds to the magnitude of the Hamiltonian and $C = 0$ for time-open transfers. Also

$$du = \frac{e\cos f df}{\ell^{1/2}} \tag{A4}$$

$$d\omega = -\frac{2(1+e\cos f)}{\ell^{1/2}} e\sin f df \tag{A5}$$

$$dr = \frac{\ell e\sin f}{(1+e\cos f)^2} df \tag{A6}$$

and

$$f = \theta - \tilde{\omega} \tag{A7}$$

where $\tilde{\omega}$ is the longitude of periapse.

$$df = d\theta - d\tilde{\omega} \tag{A8}$$

Equation (A3) becomes

$$\lambda_u \frac{e\cos(\theta-\tilde{\omega})}{\ell^{1/2}} - \lambda_\omega \frac{2e\sin(\theta-\tilde{\omega})}{\ell^{1/2}} + \lambda_r \frac{\ell e\sin(\theta-\tilde{\omega})}{[1+e\cos(\theta-\tilde{\omega})]^2}(d\theta - d\tilde{\omega}) + \lambda_\theta d\theta\Big|_{t_0}^{t_f} = 0 \tag{A9}$$

If $\tilde{\omega}_f$ and θ_f are independent, then $\lambda_\theta = 0$ and Eq. (A9) becomes

$$\lambda_u \frac{e\cos(\theta-\tilde{\omega})}{\ell^{1/2}} - \lambda_\omega \frac{2e\sin(\theta-\tilde{\omega})}{\ell^{1/2}} + \lambda_r \frac{\ell e\sin(\theta-\tilde{\omega})}{[1+e\cos(\theta-\tilde{\omega})]^2}(d\theta - d\tilde{\omega})\Big|_{t_0}^{t_f} = 0 \tag{A10}$$

where, arbitrarily, $\tilde{\omega}(t_0) = 0$. Also $df(t_0)$ and $df(t_f)$ are independent. Thus, two equations result:

$$\lambda_u \frac{e\cos\theta}{\ell^{1/2}} - \lambda_\omega \frac{2e\sin\theta}{\ell^{1/2}} + \lambda_r \frac{\ell e\sin\theta}{(1+e\cos\theta)^2}\Big|_{t=t_0} = 0 \tag{A11}$$

and

$$\lambda_u \frac{e\cos(\theta-\tilde{\omega})}{\ell^{1/2}} - \lambda_\omega \frac{2e\sin(\theta-\tilde{\omega})}{\ell^{1/2}} + \lambda_r \frac{\ell e\sin(\theta-\tilde{\omega})}{[1+e\cos(\theta-\tilde{\omega})]^2}\Big|_{t=t_f} = 0 \tag{A12}$$

These relations are automatically satisfied on a circle ($e = 0$). For $e \neq 0$, λ_θ = constant over the entire transfer and since λ_θ is proportional to A, one obtains the following theorem: Any angle-open, optimum trajectory which contains a finite circular subarc as a terminal arc or as an internal arc must also be time open. The converse of this statement is also true.

Consider the transversality equation (A12) written across an impulse where $\tilde{\omega} = 0$. The result is

$$Be^- e^+ [\sin\theta\cos(\theta-\tilde{\omega}) - 2\cos\theta\sin(\theta-\tilde{\omega}) + \sqrt{\ell^-/\ell^+}\cos\theta\sin(\theta-\tilde{\omega})] + 2e^+\sin(\theta-\tilde{\omega}) + (\sqrt{\ell^-/\ell^+} - 1)\left(B + \frac{D}{1+e^-\cos\theta}\right) = 0 \tag{A13}$$

In general, $\ell^- \neq \ell^+$ and it is required that

$$\sin(\theta-\tilde{\omega}) = 0 \tag{A14}$$

$$\sin\theta\cos(\theta-\tilde{\omega}) + (\sqrt{\ell^-/\ell^+} - 2)\cos\theta\sin(\theta-\tilde{\omega}) = 0 \tag{A15}$$

This implies,

$$\theta = n\pi \qquad n = 0,1,2,\ldots \tag{A16}$$

$$\tilde{\omega} = m\pi \qquad m = 0,1,2,\ldots \tag{A17}$$

Hence, the optimum orientation of initial, final, and intermediate orbits is coaxial and impulses will occur only at an apse. This is shown above for a junction between orbits of $e \neq 0$ and is proven by Lawden[15] for the case when one of the orbits is circular. The text clearly shows that $\tilde{\omega} = \pi, 3\pi, 5\pi, \ldots$ is nonoptimal so that orbits must not only be coaxial but the periapses must align to one side of the focus only.

References

[1] Hohmann, W., "Die erreichbarkeit der himmelskorper," Oldenbourg, Munich, 1925; also, NASA Translation TT-F-44, 1960.

[2] Shternfeld, A., *Soviet Space Science,* Basic Books, Inc., New York, 1959, pp. 109-111.

[3] Hoelker, R.E. and Silber, R., "The Bi-Elliptic Transfer Between Circular Coplanar Orbits," Department of the Army, Army Ballistic Missile Agency, Redstone Arsenal, Ala. Tech Memo 2-59, 1959.

[4] Edelbaum, T.N., "Some Extensions of the Hohmann Transfer Maneuver," *ARS Journal,* Vol. 29, 1959, pp. 864-865.

[5] Ting, L., "Optimum Orbital Transfer by Impulses," *ARS Journal,* Vol. 30, Nov. 1960, pp. 1013-1018.

[6] Ting, L., "Optimum Orbital Transfer by Several Impulses," *Astronautica Acta,* Vol. VI/Fasc. 5, 1960, pp. 256-265.

[7] Barrar, R.B., "An Analytic Proof that the Hohmann-type Transfer is the True Minimum Two-Impulse Transfer," *Astronautica Acta,* Vol. IX/Fasc. 1, 1963, pp 1-11.

[8] McIntyre, J.E., "Optimal Transfer of a Thrust Limited Vehicle between Coplanar Circular Orbits," Ph.D thesis, Princeton University, Princeton, N.J., 1964.

[9] Breakwell, J.V., "Minimum Impulse Transfer," AIAA Preprint 63-416, Aug. 1963.

[10] Contensou, P., "Etude theorique des trajectoires optimales dans un champ de gravitation. Application au cas d'un centre d'attraction unique," *Astronautica, Acta,* Vol. 8, 1962, pp. 134-150.

[11] Moyer, H.G., "Minimum Impulse Coplanar Circle-Ellipse Transfers, *AIAA Journal,* Vol. 3, April 1965, pp. 723-726.

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

[13] Hazelrigg, G.A., "The Use of Green's Theorem to Find Globally Optimal Solutions to a Class of Impulsive Transfers," AIAA Paper 68-092, Jackson Hole, Wyo., Sept. 1968.

[14] Lee, E.B., and Markus, L., *Foundations of Optimal Control Theory,* John Wiley and Sons, Inc., New York, 1967, Chap. 7.

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

[16] Pines, S., "Constants of the Motion for Optimum Thrust Trajectories in a Central Force Field," *AIAA Journal,* Vol. 2, pp. 2010-2014.

[17] Edelbaum, T.N., "How Many Impulses?" *Astronautics & Aeronautics,* Nov. 1967, pp. 64-69.

[18] Lawden, D.F., "Optimal Powered Arcs in an Inverse Square Law Field," *ARS Journal,* Vol. 31, 1961, pp. 566-568.

[19] Lawden, D.F., "Optimal Intermediate-Thrust Arcs in a Gravitational Field," *Astronautica Acta,* Vol. 8, 1962, pp. 106-123.

[20] Kopp, R.E. and Moyer, H.G., "Necessary Conditions for Singular Extremals," *AIAA Journal,* Vol. 3, 1965, pp. 1439-1444.

[21] Robbins, H.M., "Optimality of Intermediate Thrust Arcs of Rocket Trajectories," *AIAA Journal,* Vol. 3, 1965, pp. 1094-1098.

[22] Kelley, H.J., Kopp, R.E. and Moyer, H.G., *Singular Extremals,* Chap. 3, Topics in Optimization, edited by G. Leitmann, Academic Press, New York, 1967.

[23] Gurley, J.G., "Optimal-Thrust Trajectories in an Arbitrary Gravitational Field," *SIAM Journal on Control,* Vol. 2, 1964, pp. 423-432.

[24] Goh, B.S., "The Second Variation for the Singular Bolza Problem, *SIAM Journal on Control,* Vol. 4, 1966, pp. 309-325.

[25] Handelsman, M., "Aerospace System Optimization," Notes from a seminar series given at Princeton University, 1964.

[26] Miele, A., *Extremization of Linear Integrals by Green's Theorem,* Chapter 3, Optimization Techniques, edited by G. Leitmann, Academic Press, New York, 1962.

[27] Escobal, P.R., *Methods of Orbit Determination,* John Wiley and Sons, Inc., New York, 1965, Chap. 3.

[28] Ehricker, K.A., *Space Flight, Vol. I, Environment and Celestial Mechanics,* D. van Nostrand Inc., Princeton, N.J., 1960, Chap. 4.