Globally Optimal Coplanar Orbit Transfer

Huntington W. Small*
Lockheed Missiles and Space Company, Sunnyvale, Calif.

This paper considers the problem of determining which time-free transfer of a rocket between fixed elliptical orbits in a central force field requires absolutely minimum ΔV. The paper describes the operation of a program designed to determine all globally optimal (multi-impulse) solutions by joining extremal fields in state space, and it illustrates the program results by completing the solution to the coplanar version of the minimum ΔV problem. The coplanar extremals are presented both in parameter space and in state space. The advantages of parameter space are illustrated by showing how the difficult problem of "how many successive impulses can be joined in optimum transfer" can be handled there. The results presented in state space illustrate all the regions of different types of optimal transfers and map the manner in which the regions join one another. The value of this extremal-generation method is demonstrated by finding that, even in this extensively analyzed coplanar problem, the complete solution still yields some new transfer regions and some previously unguessed results; e.g., that some one-impulse transfers which continuously maximize Lawden's primer vector can be improved by finite multi-impulse transfers.

I. Introduction

THE absolutely minimum ΔV orbit transfer of a rocket between arbitrary elliptical orbits in a central force field is realized when the transfer is optimized without placing constraints on the transfer time or on the number of firings allowed. Orbit transfers of this type (which are minimum-fuel transfers when the rocket engine has a constant exit velocity) where originally investigated by Lawden in the early 1950's, and now an extensive literature has grown up around what has become known as "Lawden's Problem." It now has been established that these globally optimum transfers are composed of impulsive firings separated by coasting arcs but, despite considerable work by many analysts,[1] the location and even the number of optimal impulses have been established only in special cases. Nearly all of the 316 papers surveyed in Ref. 1 avoided the subject of global optimality, either by limiting their studies to two impulses if the impulses were to be fully optimized or by limiting the class of impulse used (e.g., assuming them to be tangent or horizontal or infinitesimal or separated by 180° coasts). Before the middle 1960's the only globally optimal orbit transfers between ellipses that had been established were the "Hohmann transfers" between coplanar coapsidal ellipses.

The subject of global optimality was revitalized in the middle 1960's by papers by Marchal,[2-4] Moyer,[5] and Winn.[6] Marchal's work on Lawden's Problem included detailed conclusions about global optimality in the coplanar version of the problem. Marchal had constructed series expansions in neighborhoods about each of the known analytical solutions to special coplanar problems and, by joining extrapolations of those results, he concluded that for coplanar time-free orbit transfer:

1) One-impulse extremals which do not extend beyond the switching conditions are always optimal if they are not improved by bi-parabolic transfers.

2) Optimal three-impulse transfers are encountered only when the following (not sufficient) conditions are met:

a) $$\left(\frac{\sqrt{P_1}}{a_1}+\frac{\sqrt{P_2}}{a_2}\right)\cdot\max(\sqrt{P_1},\sqrt{P_2})<\frac{11-\sqrt{41}}{16}$$

(which includes $\epsilon_1+\epsilon_2>(21+\sqrt{41})/16=1.7127..$);

b) $$0<|\Delta\omega|<22^\circ;\quad (18.8^\circ<|\Delta\omega|_{\max}<22^\circ)$$

c) $$\frac{9}{25}<\frac{P_1}{P_2}<\frac{25}{9}$$

d) $$\max(A_1, A_2)/\min(P_1, P_2)>21$$

where subscripted a, A, and P are terminal values of semimajor axis, apogee radius, and perigee radius in accord with Marchal's notation.

3) Finite four-impulse transfers are never optimal.

The papers by Moyer and Winn established the globally optimum solutions to transfers between coapsidal end states using the method first demonstrated by Breakwell.[7] That method is to generate the field of extremals in an orbit element state space, starting with initial state and (guessed) adjoint variables and integrating both sets while using firings which continuously maximize the instantaneous Hamiltonian. By varying the adjoints over all allowable values, and ending each extremal only when it intersects another having lower cost (which may be after many impulses), eventually state space is filled with globally optimum N-impulse trajectories from the initial orbit to every other orbit.

The author's own orbit transfer work has centered on developing and using a program which extends the extremal-field method to transfers between arbitrary ellipses in three dimensions.[8,9] In the course of generating some coplanar transfers it was found that the form of the coplanar switching surface[8] was such that Marchal's conclusions 1 and 3 could not both be true. That indicated that his series expansions had diverged from the globally optimal solutions and, so again, left wide open the question of whether more than three impulses are ever optimal. A portion of Ref. 9 then was devoted to completing the globally optimal solution to the problem of coplanar time-free orbit transfer. This paper describes the results of that study.

II. The Switching Surface in Parameter Space

It has been shown[3] that it is nonoptimal to fire out of the orbit plane in the coplanar transfer problem, so, in the notation of Ref. 8, $\sin\Upsilon=j=0$. The equations governing the generation of coplanar extremals then are expressed as functions of the four parameters $\epsilon\sin f$, $\psi\equiv 1+\epsilon\cos f$, ϕ, and $k\cos\Upsilon$. There, ϵ is the orbit's eccentricity, f is the true

Presented as Paper 76-793 at the AIAA/AAS Astrodynamics Conference, San Diego, Calif., Aug. 18-20, 1976; submitted Oct. 15, 1976; revision received June 8, 1977.

Index categories: Spacecraft Navigation, Guidance, and Flight-Path Control; Analytical and Numerical Methods.

*Staff Engineer. Member AIAA.

anomaly of the firing point, $\sin\phi$ is the radial component of the firing direction, and $k\cos\Upsilon$ is a function of adjoint variables. (The firing angles were defined so that $\cos\phi>0$. The direction of firing is set by $\cos\Upsilon$ which is $+1$ for accelerations and -1 for decelerations.) The condition for a maximum Hamiltonian requires the continuous satisfaction of the condition for a stationary Hamiltonian,

$$\tan\phi\cos\Upsilon=\frac{k\cos\Upsilon\,\epsilon\sin f}{1+\psi k\cos\Upsilon} \tag{1}$$

and the maximizing inequality:

$$\alpha_0+\alpha_1 t+\alpha_2 t^2+\alpha_3 t^3+\alpha_4 t^4\geq 0$$

for all t, where $t\equiv\tan(\Delta f/2)$ and

$$\alpha_0=\psi[1-\psi k^2-(2+\psi k^2)\tan^2\phi]$$
$$\alpha_1=4\psi\tan\phi(2k-\cos\Upsilon)-2\epsilon\sin f(1-2\tan^2\phi)$$
$$\alpha_2=\alpha_0+\alpha_4+8\epsilon\sin f\tan\phi(\cos\Upsilon-k)$$
$$-4(\psi-1)[(\cos\Upsilon-k)^2-\tan^2\phi]$$
$$\alpha_3=\alpha_1+8(\psi-1)\tan\phi(\cos\Upsilon-k)$$
$$+4\epsilon\sin f[(\cos\Upsilon-k)^2-\tan^2\phi]$$
$$\alpha_4=[(3-\psi)\cos\Upsilon-2k][2k-\cos\Upsilon]$$

Using the abbreviations,

$$z_1\equiv 4\alpha_0\alpha_4-\alpha_1\alpha_3$$
$$z_2\equiv\alpha_1{}^2\alpha_4+\alpha_0\alpha_3^2-\alpha_1\alpha_2\alpha_3$$
$$z_3\equiv 3\alpha_1^2-8\alpha_0\alpha_2$$

the maximizing inequality was shown in Ref. 8 to require

$$\alpha_0>0 \tag{2a}$$

$$16\,\alpha_0^2(\alpha_2^2+3z_1)\geq z_3 \quad \text{if} \quad z_3>0 \tag{2b}$$

$$z_1[(\alpha_2^2-z_1)^2+9\,\alpha_2 z_2]-\alpha_2^3 z_2-27\,z_2^2/4>0 \tag{2c}$$

Equation (2c) would be violated first as the α's change during a firing. Many other simpler side conditions, necessary for a maximum Hamiltonian, also could be stated, though all are included in (2); e.g., α_4 must be ≥ 0, in order that the quartic be maximized for large t.

When $\epsilon\sin f$ is eliminated from the α's with Eq. (1), then Eq. (2) becomes a function of $k\cos\Upsilon$, ψ, and $\tan^2\phi$. The $\alpha_4\geq 0$ condition is

$$1\leq 2k\cos\Upsilon\leq 3-\psi \tag{3}$$

so $k\cos\Upsilon$ is positive. The quantities $k^2\psi$, ψ, and $|\phi|$ will be taken as coordinates in what will be called "parameter space." Then Eq. (2) describes a volume in parameter space whose surface, the "switching surface," is given by the satisfaction of the equality in Eq. (2c). That is plotted in Fig. 1. The outer boundary is

$$\psi\leq 4\,k^2\psi\leq\psi(3-\psi)^2$$

from Eq. (3), and the surface represents the maximum allowable $|\phi|$ for each ψ and $k^2\psi$. (The maximum $|\phi|$ for any coplanar transfer is 26.59° at $\psi=0.355$, and $k^2\psi=0.371$.)

The surface points represents double or triple maxima of the Hamiltonian. The triple maxima occur along the dotted lines in Fig. 1. There, the surface normals are discontinuous. Each surface point not on a dotted line, then, is associated with one other surface point not on a dotted line, and each

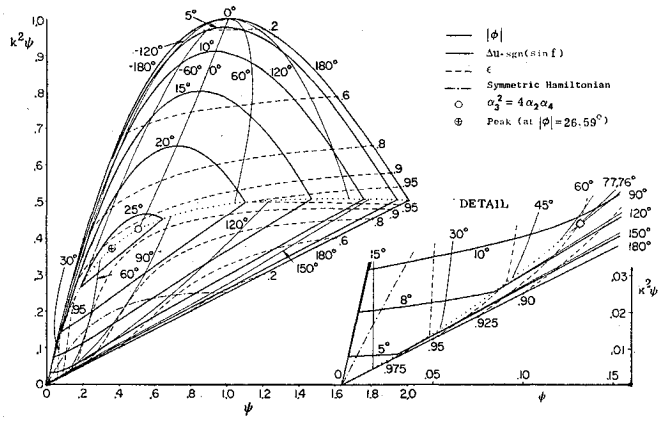

Fig. 1 The coplanar switching surface. Points $\{k^2\psi, \psi, \phi\}$ which lie within the volume maximize the Hamiltonian. Points on the surface are conditions for which the Hamiltonian has a double maximum or, along dotted lines, a triple maximum.

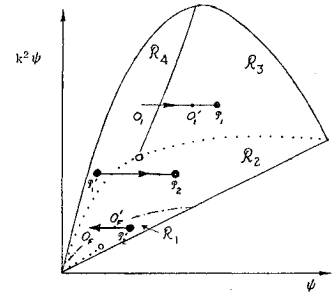

Fig. 2 Typical three-impulse switching. Switching is between "associated" points in $\mathcal{R}_1$ and $\mathcal{R}_2$ or in $\mathcal{R}_3$ and $\mathcal{R}_4$.

point on a dotted line is associated with two others on dotted lines. These associations are obtained by computing t from Eq. (A9) of Ref. 8 and then the values of the "switched" parameters at $f+\Delta f$. The values of ϵ, $|\phi|$ and $\Delta f\,\mathrm{sgn}(\sin f)$† on the surface are included in Fig. 1.

It is found that the "associated" points are coincident along both the $\Delta f=0$ line and along the line of alternating dots and dashes shown in Fig. 1. These lines represent the special cases of the Hamiltonian which are associated with the "Lawden Spiral" and the "Symmetric Transfer."[1]

These two lines and the central triple-maximum line divide the switching surface into the four regions $\mathcal{R}_1$, $\mathcal{R}_2$, $\mathcal{R}_3$, and $\mathcal{R}_4$ shown (slightly distorted) in Fig. 2. The associated points which are not coincident are found to be those pairs for which ϵ and $|\Delta f\cdot\mathrm{sgn}(\sin f)|$ are equal in $\mathcal{R}_1$ and $\mathcal{R}_2$ and in $\mathcal{R}_3$ and $\mathcal{R}_4$. The triple-maximum points have the two associations near those of neighboring points. (These triple-maximum lines will be referred to frequently in the following. Henceforth, the portions of these lines to the left and right of $\Delta f=0$ will be called 3_L and 3_R and the portion near the bottom edge of Fig. 1 will be called 3_B.)

Most of the switching surface has been investigated thoroughly in previous papers. Taken together, the only significant omission is the left central section of Fig. 1 and the region below it where the double root on the $\Delta f=0$ line interacts with the triple-maximum to make the switching sensitive.

Previous numerical studies have been discussed in Refs. 10-13. Reference 10 presents the switching in a form equivalent to Fig. 1 but just for $\epsilon=0.25$, 0.50, and 0.75. That paper also includes a brief description of the switching for $\epsilon>0.925$ by illustrating the maximum and minimum f used as a function of ϵ. The other numerical studies used formulations which encountered intrinsic difficulties as the higher values of ϵ were approached. These are described variously as:

†That is because t is proportional to $\tan\phi\cos\Upsilon$ which, from Eq. (1), depends on $\mathrm{sgn}(\sin f)$.

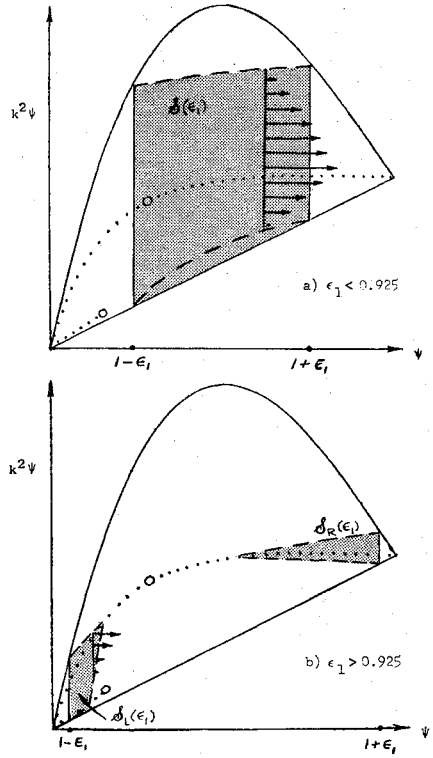

Fig. 3 The first impulse of one extremal family for a) $\epsilon_1 < 0.925$; b) $\epsilon_1 > 0.925$.

"...before the end of the acceptable range of A values was reached, convergence to the desired accuracy was not achievable by the method being employed."[11]

"These results extend only slightly beyond an eccentricity of 0.9, where the metric tensor degenerates severely, taking a different type topology."[12]

"As eccentricity increases beyond 0.8 the shape of the spool becomes 'less symmetrical' and considerably more complicated.... It is also possible to observe that the spool becomes very complicated for high eccentricities. Careful inspection allows one to note several regions where the surface intersects itself."[13]

Evidently these difficulties are all related to the appearance of the triple maximum and, since that occurrence does not limit the use of Eq. (2), that suggests that Eq. (2) will be particularly valuable for noncoplanar transfer where triple maxima occur at all eccentricities.

An input set $\{\epsilon, f, k^2\psi\}_1$ specifies ϕ from Eq. (1) and, hence, specifies a point in parameter space. Usable input to the coplanar problem is any point under the surface in Fig. 1. Since Eq. (1) is satisfied all along the extremal, the changing values of $\{\epsilon, f, k^2\psi\}$ during an impulse specify a moving point in parameter space. Since $k^2\psi$ and ϕ are constant during an impulse,[8] the paths in parameter space are straight lines parallel to the ψ axis. Since ψ/h^2 is constant during an impulse, the change of ψ is monotonic with the sign of $\cos\Upsilon$, and that is specified by the input sign of $k\psi^{1/2}$.

During the first impulse, then, an initially specified point moves as shown in Fig. 2 until it intersects the switching surface at some $\mathcal{P}_1$. The magnitude of the ΔV required is determined by the change in ψ up to its value at $\mathcal{P}_1$. The switching through Δf shifts the point in parameter space to $\mathcal{P}_1'$, the surface point associated with $\mathcal{P}_1$, and then the second impulse moves the point from $\mathcal{P}_1'$ across to another intersection at $\mathcal{P}_2$, etc. It may appear at first glance that a path which intersects the surface at one of the triple-maximum lines has an ambiguous future, but, by visualizing that extremal's neighbors, it is apparent that one choice leads to an infinitesimal impulse followed by a switch to the other choice. It is evident from Fig. 1 that switches on the $\mathcal{R}_1+\mathcal{R}_2$ surface

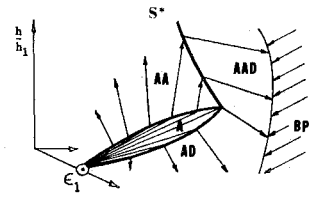

Fig. 4 Normal joining of optimal fields in state space.

are preceded by an acceleration A and followed by a deceleration D, whereas switches on the $\mathcal{R}_3$ and $\mathcal{R}_4$ surfaces separate AA or DD firings.

Although our use of Fig. 1 will be chiefly qualitative, it should be noticed that Fig. 1 yields the same information as the extremal-generation program and, if Fig. 1 were drawn accurately enough, could be used to replace it. That is, Fig. 1 completely defines the multi-impulse history of an extremal starting from any input and, with Eq. (1), continuously specifies the orbit elements and characteristic velocity along the extremal.

III. State Space

Henceforth the "cost" of an orbit transfer will refer to the dimensionless quantity $h_1\Delta V/\mu \equiv \tau$, in which ΔV is the characteristic velocity of the transfer, h_1 is the angular momentum per unit mass of the first orbit, and μ is the gravitational constant of the central field. The cost of any coplanar transfer is a function of the orbits' eccentricities (ϵ_1 and ϵ), the angle between them ($\Delta\omega$), and the orbits' angular momentum ratio. The coplanar transfer regions in state space will be described by surfaces in a set of three-dimensional state figures whose variables are ϵ, $\Delta\omega$, and h/h_1. Each state figure presents the transfers from a particular ϵ_1 to all other states by illustrating the boundary surfaces between the various transfer modes. The ϵ_1 values used will be spaced to illustrate the gradual changes in the boundaries of each of the transfer regions used in coplanar transfer.

In terms of ϵ_1 and the parameter-space variables, Eq. (1) can be written:

$$\tan^2\phi = \frac{\psi(2-\psi)-1+\epsilon_1^2}{[\sqrt{(\psi/k^2\psi)}+\psi]^2} \tag{4}$$

The points which are on the surface $\mathcal{S}(\epsilon_1)$ described by Eq. (4) and which lie under the switching surface in Fig. 1 constatute all of the input for the state figure corresponding to ϵ_1.

$\mathcal{S}(\epsilon_1)$ intersects $\phi=0$ along $\psi=1+\epsilon_1$ and $\psi=1-\epsilon_1$; it is approximately a cylindrical shell with its axis parallel to the $k^2\psi$ axis. The shells for various ϵ_1 sheath one another and intersect the switching surface along the lines shown dashed in Figs. 1, 3a, and 3b. Henceforth, those intersections with $\mathcal{R}_i$ will be called Γ_i. Shells having $\epsilon_1 > 0.925$ bulge above the central ridge of the switching surface so that their usable portions are divided into left and right parts, $\mathcal{S}_L(\epsilon_1)$ and $\mathcal{S}_R(\epsilon_1)$. Examples of the usable input regions for two state figures are illustrated as the shaded areas in Figs. 3a and 3b.

Typical input to generate extremals in a state figure is ϵ_1, f_1, and a series of $k\psi^{1/2}$ values between the allowable limits. The first impulses of two such families are shown in Figs. 3a and 3b. Each family forms a sheet in a state figure, and the set of sheets for spaced values of f_1 form layers filling the state figure for that ϵ_1. If the state figures were plotted with coordinates measuring $\epsilon\cos\Delta\omega$, $\epsilon\sin\Delta\omega$, and h/h_1, then all extremals would emanate from $\{\epsilon_1, 0, 1\}$ and the extremal family whose first impulse is illustrated in Fig. 3a would fill a sheet in state space somewhat like that illustrated in Fig. 4. The members of the family switching on $\mathcal{R}_3$ continue upward after the switch forming a sheet of extremals in the AA regions, while those members that switched on the $\mathcal{R}_2$ surface continue downward after that first switch forming a sheet of extremals in the AD region. That family and those with neighboring f will fill out portions of the A, AA, and AD regions of the state figure and, if not first improved by bi-

parabolic (*BP*) transfers, may continue on as shown in Fig. 4 to generate optimum three-impulse regions.

The region of state space which needs to be considered can be limited by the following considerations. 1) Because of the bi-parabolic transfer possibilities, it is known that no portion of any extremal will extend beyond the $\epsilon = 1$ cylinder. 2) Only the $h \geq h_1$ portion of state space need be investigated, since it is always possible to assume that the "initial" orbit is the one with the smaller angular momentum. It can be seen from Fig. 1 that, after h (and hence ψ) once decreases during a firing, it will continue to decrease on each successive firing, hence each extremal can be terminated when it first crosses the $h = h_1$ plane. It follows that only initial values of $\cos\Upsilon = +1$ are of interest here, and that requires positive initial values of $k\psi^{1/2}$ in order to satisfy Eq. (1). 3) Only the $\sin\Delta\omega > 0$ portion of state space need be considered due to symmetry. Choosing the $\sin\Delta\omega \geq 0$ half of state space to be of interest, the initial value of $\sin\phi$ (hence $d\omega/d\tau$) will be taken ≥ 0. (That is, since the initial $\cos\Upsilon$ is now $+1$, f_1 will be taken between 0 and π.)

IV. How Many Impulses?

It is useful to know in advance the number of impulses which would be optimal for each extremal before beginning to draw state figures, not only because it saves some computation time but, more importantly, it simplifies understanding the fields which appear. It can be seen immediately from Fig. 1 that the number of optimum impulses must be finite because infinitesimal impulses only occur: 1) near 3_L, in which case any one infinitesimal impulse is preceded and followed by a series of finite impulses, or 2) near the $\Delta f = 0$ line, in which case the extremal history is, indeed, a series of infinitesimal impulses (forming "stitches" along the $\Delta f = 0$ line), but in the limit those generate the nonoptimal Lawden Spiral. Marchal has stated that even the possibility of a fourth-impulse can be disregarded in any coplanar transfer, but those statements need further consideration.

Consider the history of that portion of the ϵ_1, f_1 family in Fig. 3a which switches near the triple-maximum. The members switching on $\mathcal{R}_3$ switch again on $\mathcal{R}_2$ after a small acceleration and then decelerate, which those members switching on $\mathcal{R}_2$ make neighboring decelerations immediately as sketched in Fig. 4. (If the initial ψ were smaller so that the first impulse were larger, then the *BP* region in Fig. 4 might extend into the one-impulse region and eliminate the *AAD* region.) Figure 4 is the typical joining of *A*, *AA*, *AD*, *AAD*, and *BP* regions using switching surfaces near to $\psi = 0$ and $\psi = 2$. It corresponds to Marchal's solution for transfers using coasting arcs with $\epsilon \to 1$. Now consider an extremal family which begins just to the right of some point $\mathcal{P}_0$ on $\mathcal{R}_4$ and switches near the triple maximum at large ϕ. The first impulse is small so that the *BP* region is not near the multiple joining of the fields. Then, if the *AAD* field were still optimal, the four-impulse extremals whose first impulses were small and switched near the point associated with $\mathcal{P}_0$ also would be optimal. If four impulses are never optimal, then the region of state space filled by the *AAD* maneuver must be filled by some other (non-neighboring, non-*BP*) field forming a comparison surface with the *AA*, *A*, and *AD* fields (and contradicting Marchal's conclusion 1). If no such new comparison field exists, then an optimum four-impulse field exists.

To investigate this situation, some families of extremals starting near the triple maximum near maximum ϕ were generated. It was found that the S^* switching surface in Fig. 4 now bulged out over the *A-AA* switching surface, *reflecting* the *AA* field near the tip of the *A* field and so leaving an empty space (empty of any neighboring extremal field) to the right of the tip. Then evidently some comparison field must be improving the whole region of the tip. Extending the range of the input adjoints being tested, it was found that for high-ϕ input the triple-maximum fields join as pictured in Fig. 5. The missing field filling the region to the right of the tip is still an *AAD* field, but now the extremals are those which switched

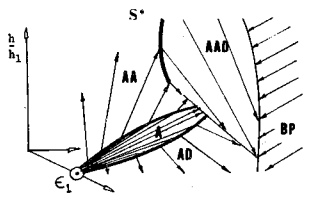

Fig. 5 Abnormal joining of optimal fields in state space.

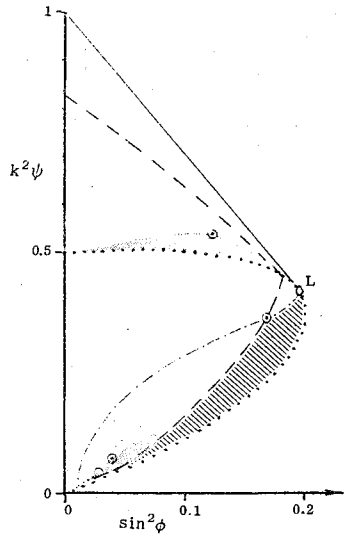

Fig. 6 Optimal three-impulse limits.

much higher up on S^* where S^* was still passing extremals. The usable portion of S^* extends only down to where its curvature is so great that the *AAD* extremals leave S^* tangent to it; below that tangency line the *AAD* extremals form comparison surfaces with the *AA*, *A*, and *AD* regions.

That sensitive high-ϕ region also can be reached by extremals which first switch near 3_B, so some additional extremal families were initiated near $\psi = 0$. Again it was found that part of the second switching surface sometimes reflected extremals, but, because that part did not overhang the tip of the one-impulse surface, the comparison field which improved the region of the reflecting surface did not improve any one-impulse extremals. However, those fields did include optimal three-impulse transfers which had perigee ratios up to about 20% greater than Marchal's limit of 25/9.

Details of the fields just discussed will be illustrated in the following sections. Here we are concerned with the fact that Marchal's results are not applicable to that sensitive region in which bi-parabolic cost does not limit immediately multi-impulse possibilities, and (because there may well be more new fields than the trial runs discussed uncovered) we will devote the rest of this section to reconsidering what input if any could possibly lead to an optimum four-impulse coplanar transfer.

Figure 2 illustrates a typical three-impulse extremal in parameter space. It extends from 0_I to 0_F and passes through points $0_I'$ and $0_F'$. The points $0_I'$ and $0_F'$ specify the eccentricities there, and the path between them specifies a change in ω and an angular momentum ratio. If the $0_I' - 0_F'$ transfer is not optimal, then the $0_I - 0_F$ transfer will not be optimal. By starting extremals at all points $0_I'$ near the first switching surface and testing $\tau_{BP} - \tau$ and D for a change of sign before (or at) the second switch,‡ it is possible to determine which surface points $\mathcal{P}_1$ can be used for the first switch of optimal three-impulse transfers. It is sufficient to

‡D is the determinant formed of the partial derivatives of the state with respect to input parameters and cost. D changes sign continuously at a conjugate point and discontinuously at a "switch-conjugate" point (where a switching surface reflects extremals back into their own field.)[9]

test points on $\mathcal{R}_3$. Starting from points which are inside $\mathcal{R}_3$ by some small constant $\Delta\psi$, the values of $\tau_{BP}-\tau$ and D at the second switch are just piecewise continuous functions of $\mathcal{P}_1$, and there is no difficulty encountered in using a reasonable number of runs to determine the regions on $\mathcal{R}_3$ where these functions are positive or negative.

The results are plotted in Fig. 6, which shows a "side view" of Fig. 1 so that all extremals are represented as points. To the right of the dashed line, which extends down to $k^2\psi=0$, the second impulse costs less than the bi-parabolic maneuver. To the right of the line of alternating dots and dashes, which extends down to a point on the lower triple maximum, the second impulse is not improved by a neighboring extremal. Then any optimum three-impulse maneuver's second impulse must be in the hatched area, and that requires its first (and third) impulse to be in one of the shaded areas.

In order to have an optimal four-impulse maneuver, both the second and third impulses must lie in the hatched area. That can occur only when a shaded region covers some of the hatching. The regions touch at L, (where the $\Delta f=0$ line ends), but have no common area. The lower shaded area does include a narrow piece below the dashed line, but there each full impulse (from surface to surface) has approximately bi-parabolic cost so the sum of the costs is far from optimum, and we can conclude that four-impulse transfers are never optimal for coplanar transfer.

Now consider the bounds on the three-impulse transfers. The extremals which did complete the second switch satisfactorily were continued until the test conditions first failed. The interior points of first failure form surfaces in front of each of the shaded areas shown in Fig. 6. These surfaces, with the shaded portions of the switching surfaces and the extremal paths reaching the triple maxima, form small volumes enclosing all of the input points which could ever lead to optimum three-impulse transfers. Those input volumes are plotted in detail in Ref. 9. (At this point, of course, it is not known whether an extremal generated from one of those volumes can be improved by some non-neighboring finite transfer before the second switch. It is found, however, that that does not occur, so any input within one of those volumes does generate some optimum three-impulse transfers.)

V. State Figures

One-Impulse Regions

Each state figure applies to one initial ellipse which is specified by the value of its eccentricity ϵ_1. The state figure describes the transfer from that orbit to all coplanar ellipses. For low and moderate values of ϵ_1, the geometry of a state figure is simplified by using a state vector $\bar{x}$, whose components are $(h_1/h)^2\ \epsilon\ \cos\Delta\omega$, $(h_1/h)^2\ \epsilon\ \sin\Delta\omega$, and $1-(h_1/h)^2$, measured along orthogonal axes as shown in Fig. 7. (Henceforth we will use the abbreviations: $\epsilon_\parallel \equiv \epsilon\cos\Delta\omega$, $\epsilon_\perp \equiv \epsilon\sin\Delta\omega$, and $z\equiv 1-(h_1/h)^2$.) During the first impulse, $\bar{x}$ moves along the path

$$\bar{x}=\bar{x}_1+z(\bar{e}_R+\gamma\bar{e}_L+\bar{e}_h)$$

in which $\gamma\equiv\tan\phi/(k_1+k)$ and $\bar{e}_R$, $\bar{e}_L$, $\bar{e}_h$ are orthogonal unit vectors: $\bar{e}_h$ points along the z axis; $\bar{e}_R$ and $\bar{e}_L$ make angles of f_1 and $f_1+90°$ with the $\epsilon_\parallel$ axis. (It is useful to visualize the instantaneous orbit in physicial space in the $\epsilon_\parallel$, $\epsilon_\perp$ plane with its focus at the origin and the horizontal component of $\bar{x}$ pointed toward its perigee. Then $\bar{e}_R$ is in the direction of the radius to the location of the impulse on the instantaneous orbit, and $\bar{e}_R$ rotates by Δf at switches.)

If the initial impulse had been applied tangentially, then γ would be zero, and $\bar{x}-\bar{x}_1$ would be a straight line in state space. The set of tangential firings from all f_1 on ϵ_1 then would describe the surface of a cone whose apex is at $\bar{x}_1$. Optimum transfers, however, are not ordinarily tangential, since $\tan\phi$ is only zero when $\epsilon\ \sin f=0$, nevertheless the position of the "tangent cone" in state space is still physically meaningful, for state points outside the tangent cone represent orbits which intersect the initial orbit, and state points inside represent orbits which do not.

Since the variation in $\bar{x}-\bar{x}_1$ between the members of one "extremal family" (see Figs. 3a and 3b) is caused only by γ, an extremal family fills a portion of a plane tangent to the tangent cone. Extremals from the "families" neighboring f_1 fill sections of neighboring planes which lie over one another like shingles so that the one-impulse extremals do not intersect one another. The collection of all one-impulse extremals from $\mathcal{S}(\epsilon_1)$ fills a thin conical region slightly outside (or touching) the tangent cone in the corresponding ϵ_1 state figure. The A region surrounds the tangent cone for $\epsilon_1<0.925$, but it becomes very thin near the maximum $\epsilon_\perp$ regions in those state figures having ϵ_1 slightly less than 0.925, and it divides into two one-impulse regions corresponding to f_1 families from $\mathcal{S}_L(\epsilon_1)$ and $\mathcal{S}_R(\epsilon_1)$ when $\epsilon_1>0.925$. The A regions are optimal as long as they are not improved by BP transfers or AAD comparison fields of the type shown in Fig. 5. (These AAD comparisons only occur when $0.81<\epsilon_1<0.925$; the details of which transfers are improved are graphed in Ref. 9.)

Two-Impulse Regions

The points on the lower surface of a one-impulse region are the switching states on $\mathcal{R}_1$ and $\mathcal{R}_2$. During the second impulse, extremals switching at those points continue downward and outward from the one-impulse region, forming the AD region. Those extremals usually fill the space out to the BP boundary uniquely, but there are exceptions for $\epsilon_1>0.81$ when the comparison surfaces with three-impulse regions discussed earlier first appear. Their locations will be shown later when three-impulse regions are discussed.

Those points on the upper surface of a one-impulse region at which switching occurs are the switching states reached on $\mathcal{R}_3$, but a portion of the upper surface in each state figure is composed of extremal paths. (That portion is between C_2 and C_4 in Fig. 7.) During the second impulse, the AA extremals switching on the upper surface of the one-impulse region continue upward and (generally) inward from the one-impulse region and so tend to meet in the central portions of the tangent cone. This meeting is along a line of conjugate points on the $\epsilon=0$ axis when $\epsilon_1=0$, and the meeting forms a comparison surface for higher ϵ_1. These comparison surfaces appear for all $\epsilon_1<0.925$, but most of each surface is covered by BP regions unless ϵ_1 is small. Points on the surfaces are reached with equal cost by two "equivalent" extremals, as

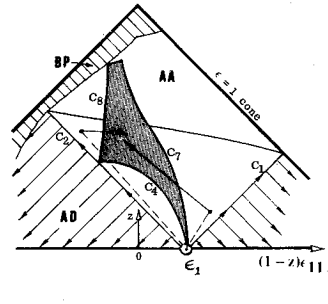

Fig. 7 Optimal transfer regions when $\epsilon_1<0.713$.

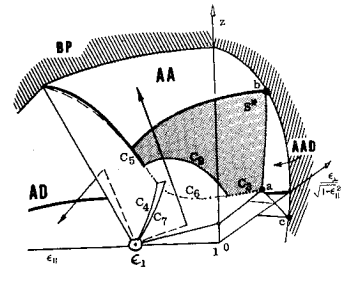

Fig. 8 Optimal transfer regions when $\epsilon_1<0.925$.

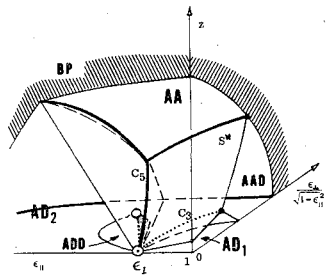

Fig. 9 Optimal transfer regions when $0.925 < \epsilon_I < 0.975$.

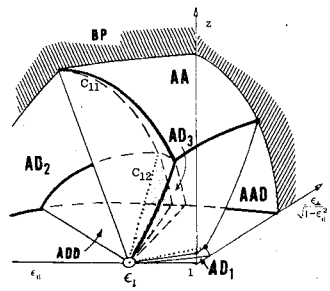

Fig. 10 Optimal transfer regions when $0.975 < \epsilon_I < 1$.

illustrated in Fig. 7. A comparison surface intersects the $\epsilon_\perp = 0$ plane along C_8. It intersects the one-impulse region along C_4, and there the "equivalent transfers" are those from associated points on Γ_3 and Γ_4. Points near C_7 are reached by neighboring extremals, and C_7 is a line of conjugate points. The value of ϵ on a comparison surface is nearly constant at any z so that they appear as arcs of circles in horizontal sections through state space. For any ϵ_I and z, the limits of $\Delta\omega$ on those arcs are found to be very nearly

$$\min\left[1, \frac{z^2(1-11\epsilon_I^2/18)}{4\epsilon_I^2(1-5z/6)}\right] \geq \sin^2\frac{\Delta\omega}{2}$$

$$\geq \frac{z^2(1-11\epsilon_I^2/18)}{4\epsilon_I^2(1-5z/6)+z^2(1+\epsilon_I^2/4)}$$

and (using $\ell \equiv h/h_I$) the eccentricity there is given to within 2% by

$$\epsilon_{AA} = \epsilon_I\left[1 + \frac{(\ell-1)(\ell^2+1)}{2\ell^2+\ell+1+\ell\sqrt{2\ell^2+2}}\right]$$

Three-Impulse Regions

The three-impulse regions shown in Figs. 4 and 5 are always a feature of switching near a triple maximum, but those regions are improved by the *BP* field for low ϵ_I. Only when $\epsilon_I > 0.713$ do optimal three-impulse regions appear in the state figures. The three-impulse regions join the one-impulse and two-impulse regions in three ways, depending on ϵ_I. Those are illustrated (slightly distorted to show more detail) in Figs. 8-10. Note that the coordinate axes there measure variables which are different from those in Fig. 7.

Input to the *AAD* region is given by the boundaries of the volume in front of the upper shaded region in Fig. 6.[9] The intersections of $\mathcal{S}(\epsilon_I)$ with the surface of that volume correspond to the boundaries of the S^* switching surface (which is shown shaded in Fig. 8, but not in Figs. 9 and 10). The S^* surface first appears when *a*, *b*, and *c* coincide.§ C_9 is that limit to S^* at which extremals leave the switching surface tangent to it. Below C_9, and extending down to $z=0$ is the *AAD* comparison surface discussed in reference to Fig. 5. C_9 and that comparison surface only occur in state figures having $0.81 < \epsilon_I < 0.925$. The extremals from input points on Γ_3 would make a zero-impulse, switch immediately to Γ_4 and then switch on C_5. When $\epsilon_I > 0.925$, C_5 extends down to $\dot{x}_I$, and then the surface formed by the Γ_3 extremals and Γ_4 extremals divides state space into two portions as shown in Fig. 9. The left portion is reached only by extremals from $\mathcal{S}_L(\epsilon_I)$ and the right portion is reached only from $\mathcal{S}_R(\epsilon_I)$. When $\epsilon_I > 0.925$, some extremals from $\mathcal{S}_L(\epsilon_I)$ contact 3_B, and then a pyramidal *ADD* region appears in the state figures. The curve C_{12} in Figs. 9 and 10 represents states reached at 3_B. It is analogous to C_3, and sections through C_{12} would exhibit fields like those shown in Fig. 4. The pairs of extremals which first switch at neighboring points just above and below 3_B continue as neighbors (in state space and parameter space) after one makes an infinitesimal impulse near 3_L. These neighbors form the left surface of the *ADD* region. The right surface, which will be called S^{**}, is the second switching surface. That is shown shaded in Figs. 9 and 10. The outermost surface of the pyramid is either a comparison surface (as in Fig. 9) or a *BP* boundary (as in Fig. 10) or a combination of both. The inner end of 3_B (and C_{12}) is not reached before *BP* cost is exceeded when $\epsilon_I > 0.975$. Then C_{12} joins C_{11} (which is composed of states at which extremals have just *BP* cost at their first switch), so a third narrow *AD* region (AD_3) appears between S^{**} and the *AAD* region. (The $\epsilon = \epsilon_I$ curve in the $z=0$ plane is near the center of the AD_3 region, so the "symmetric transfers" between highly eccentric orbits are contained in AD_3.) The optimality regions shown in Fig. 10, which are those discussed by Marchal, then remain typical while $\epsilon_I \to 1$.

§This first appearance occurs when transferring to nearly parabolic orbits. Those transfers are discussed in detail in Marchal's excellent papers.

Table 1 Samples of minimum fuel orbit transfer

Case	JMAX	$\frac{h_1}{\mu}\sum\Delta v$	$\epsilon\cos\Delta\omega$	$\epsilon\sin\Delta\omega$	$(h/h_1)^2$	f	ϕ
1	1	0.000000	0.875000	0.000000	1.000000	128.6537	25.0681
		0.085341	0.894220	0.230220	1.370000		
2	3	0.000000	0.875000	0.000000	1.000000	90.8255	20.1909
		0.038429	0.919644	0.074103	1.074390	148.6340	24.5207
		0.079502	0.893033	0.209088	1.502393	-158.1635	11.0362
		0.085070	0.894220	0.230220	1.370000		
3	2	0.000000	0.400000	0.000000	1.000000	160.5105	3.9723
		0.325092	-0.336022	0.370517	2.312317	-165.4847	2.8733
		0.348858	-0.400000	0.400000	2.000000		
4	2	0.000000	0.825000	0.000000	1.000000	120.6329	22.1953
		0.093700	0.856410	0.235346	1.321753	-157.1801	10.3767
		0.107770	0.860000	0.280000	1.100000		
5	2	0.000000	0.940000	0.000000	1.000000	165.2187	9.2171
		0.025057	0.903590	0.082218	1.616645	-89.3651	20.0922
		0.056144	0.850000	0.150000	1.500000		
6	2	0.000000	0.600000	0.000000	1.000000	43.8556	8.1258
		0.098233	0.774830	0.108288	1.140365	166.4101	10.6173
		0.297444	0.200000	0.400000	4.000000		
7	2	0.000000	0.200000	0.000000	1.000000	35.8098	2.7730
		0.222949	0.616425	0.256946	1.419934	174.2475	3.5201
		0.470220	-0.220000	0.100000	5.000000		
8	3	0.000000	0.940000	0.000000	1.000000	68.7896	16.4532
		0.010219	0.955810	0.014966	1.014680	158.6777	21.7180
		0.025174	0.945600	0.072880	1.290611	-162.2056	9.8126
		0.031826	0.950000	0.100000	1.100000		
9	3	0.000000	0.965000	0.000000	1.000000	166.0902	9.0942
		0.012303	0.950413	0.051867	1.420669	-155.7925	22.0702
		0.021390	0.952202	0.090463	1.212631	-69.3602	16.7952
		0.028013	0.940000	0.100000	1.200000		
10	2	0.000000	0.980000	0.000000	1.000000	168.2953	11.9869
		0.007377	0.972180	0.039747	1.389342	-165.9539	11.9876
		0.014753	0.976729	0.080000	1.000000		

VI. Examples

Examples of each type of globally optimal coplanar orbit transfer are presented in Table 1. The input parameters needed to generate the optimal transfers were determined iteratively. On each iteration a "nominal" and two neighboring extremals were generated simultaneously up to the specified final z on the specified final impulse. The observed changes in $\epsilon_\parallel$ and $\epsilon_\perp$ between the nominal and its neighbors

then were used to adjust the nominal input parameters toward those necessary to reach the final target state. The process was iterated until the nominal extremal's state at the final z was within 1×10^{-7} of the target state. The total computation time for all 10 cases was less than 9 sec on an IBM 360-67.

The columns in Table 1 show for each case: the specified number of impulses (JMAX), the cost and state at the beginning or end of one impulse, and the angles describing the next impulse. Briefly, the examples given in the table are as follows:

a) Cases 1 and 2 relate to Figs. 5 and 8. Case 1 is a one-impulse transfer which lies under the switching surface, and so is potentially optimal, but the final state lies outside the C_6 line for $\epsilon_1=0.875$ so it can be improved by an AAD transfer. Case 2 is the AAD transfer which improves it.

b) Cases 3, 4, and 5 are various AD transfers.

c) Cases 6 and 7 are examples of AA transfers. Case 6 is the usual case in which the final state is clearly within the tangent cone and not near the AA comparison surface. Case 7 illustrates the situation in which the final eccentricity is slightly less than the value on the AA comparison surface. (Then the final state might be reached from two points in parameter space, but the input from the smaller f_1 is optimal.)

d) Cases 8 and 9 are typical AAD and ADD transfers from $\mathcal{S}_R(\epsilon_1)$ and $\mathcal{S}_L(\epsilon_1)$, respectively.

e) Case 10 is an example of an AD_3 transfer. Note that this case is nearly a "symmetric transfer."

VII. Conclusions

This paper has illustrated the use of an efficient program which generates minimum-fuel extremals. It has demonstrated its value by completing the globally optimal solution to one of the most extensively analyzed problems in orbit optimization theory.

Often the program has been used to generate (locally) optimal transfers between noncoplanar ellipses using approximate analytical results[1] as first estimates of the optimal, but no attempt has been made to complete the general noncoplanar transfer problem in the manner that the coplanar problem was treated here. The general noncoplanar transfer between ellipses involves two more parameters than the coplanar problem (and, in addition, it is known that triple maxima occur at all eccentricities in noncoplanar transfers, so one might expect that each of the many projections needed to describe state space would be at least as complicated as the most complicated regions of coplanar state figures), so it appears that the globally optimal solutions to that general problem will not be forthcoming in the near future. The construction of the state figures in the coplanar problem, however, required far more computation and geometric detail than did the study of parameter space, and it was seen that much information about coplanar transfers was obtained from Figs. 1 and 6. Then it seems that a valuable and relatively simple extension to the present work would be the computation of a set of figures like Figs. 1 and 6 for the range of nonzero values of $j\psi^{-1/2}$ and $\psi^{1/2}\sin\Upsilon$ (perhaps using ϕ, $k\psi^{1/2}$, $j\psi^{-1/2}$, $\psi^{1/2}\sin\Upsilon$, and $\psi^{1/2}\cos\Upsilon$ as parameters, so that, again, only one parameter varies during an impulse). It is hardly possible to anticipate all of the information to be derived from such a study, but at least it is expected that valuable conclusions concerning four-impulse transfers could be drawn and, perhaps, some interesting regions of parameter space in which the switching surfaces are qualitatively different from the coplanar switching surface would be isolated for more detailed study.

References

[1] Gobetz, F.W. and Doll, J.R., "Survey of Impulsive Trajectories," *AIAA Journal*, Vol. 7, May 1969, pp. 801-834.

[2] Marchal, C., "Transferts Optimaux Entre Orbites Elliptiques Coplanaires (Durée Indifférente)," *Astronautica Acta*, Vol. 11, Nov.-Dec. 1965, pp. 432-445.

[3] Marchal, C., "Transfers Optimaux Entre Orbites Elliptiques (Durée Indifférente)," Ph.D. Thesis, June 1967, University of Paris.

[4] Marchal, C., "Synthése des Résultats Analytiques sur les Transferts Optimaux entre Orbites Képlériennes (Durée Indifférente)," *Advanced Problems and Methods for Space Flight Optimization*, (Proceedings of a Colloquium at University of Liege, Belgium), edited by B. Fraeijs de Veubeke, Pergamon Press, New York, 1969, pp. 91-156.

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

[6] Winn, C.B., "Minimum-Fuel Transfers Between Coaxial Orbits, Both Coplanar and Noncoplanar," American Astronautical Society, Paper 66-119, July 1966.

[7] Breakwell, J.V., "Minimum Impulse Transfer," *AIAA Progress in Astronautics and Aeronautics*, Vol. 14, edited by V.G. Szebehely, Academic Press, New York, 1964, pp. 583-589.

[8] Small, H.W., "Minimum N-Impulse Time-Free Transfers Between Elliptic Orbits," *AIAA Journal*, Vol. 9, April 1971, pp. 594-599.

[9] Small, H.W., "Minimum Fuel Time-Free Transfer Between Elliptic Orbits," Ph.D. Dissertation, Aug. 1972, Stanford University.

[10] Moyer, H.G., "Necessary Conditions for Optimal Single-Impulse Transfer," *AIAA Journal*, Vol. 4, Aug. 1966, pp. 1405-1410.

[11] Gobetz, F.W., Washington, M., and Edelbaum, T.N., "Minimum Impulse Time-Free Transfer Between Elliptic Orbits," NASA CR-636, Nov. 1966.

[12] Culp, R.D., "Contensou-Busemann Conditions for Optimal Coplanar Orbit Transfer," *AIAA Journal*, Vol. 5, Feb. 1967, pp. 371-372.

[13] McCue, G.A. and Bender, D.F., "Contensou's Spool," *Journal of the Astronautical Sciences*, Vol. 14, Sept.-Oct. 1967, pp. 241-242.