Simple Proof of the Global Optimality of the Hohmann Transfer

John E. Prussing*
University of Illinois at Urbana-Champaign, Urbana, Illinois 61801

Introduction

IN 1925, Hohmann[1] conjectured that the minimum-fuel transfer between coplanar circular orbits in an inverse-square gravitational field is the two-impulse elliptic transfer that is tangent to both of the terminal circular orbits. The first impulse inserts the vehicle into the transfer orbit, while the second impulse circularizes the orbit at the final radius. Each velocity change is parallel to the instantaneous velocity vector, apparently resulting in the most fuel-efficient maneuver.

It was nearly 40 years before proofs of optimality among the class of two-impulse transfers were given by Lawden,[2] using the calculus of variations, and by Barrar.[3] Proofs were also provided by Hazelrigg[4] using Green's Theorem, Marec[5] by graphical construction, and Battin[6] using Lagrange multipliers. Most recently, Palmore[7] presented an elementary proof using ordinary calculus.

The Hohmann transfer itself is the easiest transfer orbit to calculate all of the feasible conic transfer orbits (those which intersect both circular terminal orbits and can therefore be used for a two-impulse transfer.) The velocity changes can be calculated by a scalar rather than vector velocity subtraction, the semimajor axis of the transfer ellipse is just the arithmetic mean of the circular terminal radii, and the transfer time is simply half of the period of the transfer orbit. The irony is that even though the Hohmann transfer is the easiest to calculate, in order to prove its global optimality, one must compare its fuel cost with all of the other feasible transfer orbits, which are considerably more difficult to calculate.

In Palmore's proof, each possible transfer orbit is described by two variables that are nonlinear functions of the parameter (semilatus rectum) p and the eccentricity e. The boundaries of the region of the (p,e) plane that contain feasible transfer orbits are then determined as relatively complicated curves and the gradient of the characteristic velocity of the maneuver (the sum of the velocity changes, representing the total fuel cost) with respect to these variables is calculated in order to determine the constrained minimum.

The proof that follows is similar to Palmore's, but there are two essential differences. First, the variables used are the familiar parameter and eccentricity themselves, rather than nonlinear functions of them. This results in the boundaries of the feasible region in the (p,e) plane being simple straight lines.

Second, whereas Palmore utilizes the gradient of the characteristic velocity with respect to both variables, the current proof requires only the (simpler) partial derivative of the characteristic velocity with respect to the eccentricity.

Feasible Region

To determine the boundaries of the feasible region, one utilizes the basic polar equation for a conic orbit:

$$r = p/(1 + e \cos f) \tag{1}$$

Received Jan. 29, 1991; revision received Oct. 28, 1991; accepted for publication Nov. 13, 1991.

*Professor, Department of Aeronautical and Astronautical Engineering. Associate Fellow AIAA.

Feasibility requires that the transfer orbit intersect both circular terminal orbits of radii r_1 and r_2, where $R \equiv r_2/r_1 > 1$ is assumed regardless of which are the initial and final terminal radii. Infeasible orbits are those for which either the periapse lies outside the smaller terminal orbit or the apoapse lies inside the larger terminal orbit. The conditions to be satisfied are from Eq. (1):

$$r_p = p/(1 + e) \le r_1 \tag{2}$$

$$r_a = p/(1 - e) \ge r_2 \tag{3}$$

Equations (2) and (3) can be rewritten as

$$p \le r_1(1 + e) \tag{4}$$

$$p \ge r_2(1 - e) \tag{5}$$

Each possible conic transfer orbit can be characterized by its values of p and e and represented as a point in the (p,e) plane. The inequalities (4) and (5) define the region of feasible transfers. As shown in Fig. 1, the boundaries of this region are straight lines. The region of feasible transfer orbits contains ellipses for $0 < e < 1$, parabolas for $e = 1$, and hyperbolas for $e > 1$, with all of the rectilinear conic orbits located at point R for which $e = 1$ and $p = 0$.

Hohmann Transfer Optimality

The magnitude of a velocity change to depart or enter a circular orbit at $r = r_k$ where $k = 1,2$ is described by the law of cosines as

$$(\Delta v)^2 = v^2 + v_c^2 - 2v_c v_\theta \tag{6}$$

where v_c is circular orbit speed and v_θ is the component of the velocity vector normal to the radius. Using conservation of angular momentum, $v_\theta = h/r = (\mu p)^{1/2}/r$, where h is the specific angular momentum of the orbit and μ is the gravitational parameter. In addition, from conservation of energy (the vis-viva equation),

$$v^2 = \mu[(2/r) - (1/a)] \tag{7}$$

where a is the semimajor axis, it follows that $v_c^2 = \mu/r$ and, using $p = a(1 - e^2)$,

$$v^2 = \mu\{(2/r) + [(e^2 - 1)/p]\} \tag{8}$$

The characteristic velocity is then

$$\Delta v_T = \Delta v_1 + \Delta v_2$$

where Δv_1 occurs at $r = r_1$ and Δv_2 occurs at $r = r_2$. Thus Δv in Eq. (6) can be replaced by Δv_k corresponding to $r = r_k$. If one examines the gradient of the characteristic velocity with respect to the eccentricity of the transfer orbit, one obtains

$$\partial\Delta v_T/\partial e = (\partial\Delta v_1/\partial e) + (\partial\Delta v_2/\partial e) \tag{9}$$

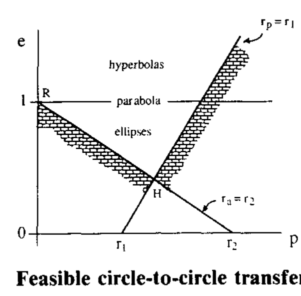

Fig. 1 Feasible circle-to-circle transfer orbits.

and, from differentiating Eq. (6) using Eq. (8),

$$2\Delta\nu_k\,(\partial\Delta\nu_k/\partial e) = 2\nu_k\,(\partial\nu_k/\partial e) = 2e\mu/p \tag{10}$$

Thus

$$\partial\Delta\nu_T/\partial e = (e\mu/p)(\Delta\nu_1^{-1} + \Delta\nu_2^{-1}) > 0 \tag{11}$$

The fact that the partial derivative of the characteristic velocity with respect to the eccentricity is positive means that at any point (p,e) in the interior of the feasible region the characteristic velocity can be decreased by decreasing the value of e while holding the value of p constant. This means that the minimum characteristic velocity will always lie on the cross-hatched boundary of the feasible region given by the union of $r_a = r_2$ and $r_p = r_1$ in Fig. 1.

What will be demonstrated is that when the characteristic velocity is constrained to the boundary, its derivative with respect to e remains positive. Thus the optimal solution lies at the minimum value of e on the boundary, i.e., at point H in Fig. 1, which represents the Hohmann transfer.

To do this, the expression for the characteristic velocity is constrained to the boundary by substituting for the value of p along each portion of the boundary, namely $p = r_2(1-e)$ on the left part $(r_a = r_2)$ and $p = r_1(1+e)$ on the right part $(r_p = r_1)$. Thus the characteristic velocity becomes a function of the single variable e. Letting $\Delta\tilde{\nu}$ denote the velocity change constrained to the boundary, Eq. (6) becomes for the left part

$$(\Delta\tilde{\nu})^2 = \mu\left[\frac{2}{r} + \frac{e^2-1}{r_2(1-e)}\right] + \frac{\mu}{r} - 2\left(\frac{\mu}{r}\right)^{1/2}\frac{[\mu r_2(1-e)]^{1/2}}{r} \tag{12}$$

Differentiating with respect to the single variable e and recalling that $R = r_2/r_1 > 1$ and $0 < e < 1$,

$$2\Delta\tilde{\nu}_1\frac{d\Delta\tilde{\nu}_1}{de} = \frac{\mu}{r_2}\left[R^{3/2}\frac{1}{(1-e)^{1/2}} - 1\right] > 0 \tag{13}$$

and

$$2\Delta\tilde{\nu}_2\frac{d\Delta\tilde{\nu}_2}{de} = \frac{\mu}{r_2}\left[\frac{1}{(1-e)^{1/2}} - 1\right] > 0 \tag{14}$$

and thus, similar to Eq. (11), $d\Delta\tilde{\nu}_T/de > 0$ on the left portion of the boundary.

Similarly, on the right portion of the boundary,

$$(\Delta\tilde{\nu})^2 = \mu\left[\frac{2}{r} + \frac{e^2-1}{r_1(1+e)}\right] + \frac{\mu}{r} - 2\left(\frac{\mu}{r}\right)^{1/2}\frac{[\mu r_1(1+e)]^{1/2}}{r} \tag{15}$$

Differentiating, one obtains

$$2\Delta\tilde{\nu}_1\frac{d\Delta\tilde{\nu}_1}{de} = \frac{\mu}{r_1}\left[1 - \frac{1}{(1+e)^{1/2}}\right] > 0 \tag{16}$$

and

$$2\Delta\tilde{\nu}_2\frac{d\Delta\tilde{\nu}_2}{de} = \frac{\mu}{r_1}\left[1 - \frac{1}{R^{3/2}(1+e)^{1/2}}\right] > 0 \tag{17}$$

and thus, $d\Delta\tilde{\nu}_T/de > 0$ on the right portion of the boundary also. Because the derivative of $\Delta\tilde{\nu}_T$ with respect to the eccentricity is positive at all points on the boundary, it follows that the optimal solution lies at the point of minimum eccentricity on the boundary, namely at point H in Fig. 1, representing the Hohmann transfer. This proves the global optimality of the Hohmann transfer among two-impulse transfers.

As mentioned by Edelbaum,[8] for values of R greater than approximately 11.94 (printed as 11.24 in Ref. 8 due to a typographical error), the Hohmann transfer is no longer the global optimal impulsive transfer. A three-impulse bielliptic transfer can always be found that has a lower cost, if the midcourse impulse occurs sufficiently far outside the outer terminal orbit. For R greater than approximately 15.58, any bielliptic transfer that has its midcourse impulse outside the outer terminal orbit has a lower cost than the Hohmann transfer.[2,6,8] The major disadvantage of the bielliptic transfer is that the transfer time is more than twice the Hohmann transfer time, while the cost saving in the coplanar case is modest. In the noncoplanar case, the bielliptic transfer provides significant cost savings[8] because the required plane change can be accomplished at larger radius and therefore lower velocity.

There exists no optimal bielliptic transfer, but the infimum of the bielliptic transfers is the biparabolic transfer, which is the limiting case of the bielliptic transfer as the midcourse impulse location tends to infinity. However, for all values of R, the Hohmann transfer is the globally optimal two-impulse transfer, based on the proof given above.

Concluding Remarks

The case of two-impulse transfer between coplanar circular orbits is investigated. By using the familiar orbital elements, eccentricity e and parameter (semilatus rectum) p, the global optimality of the Hohmann transfer among the class of two-impulse transfers is proved using ordinary calculus. The proof presented is simpler than existing proofs in the literature.

Acknowledgment

This research was supported by NASA Research Grant NAG 3-1138 administered by the Systems Analysis Branch of the NASA Lewis Research Center.

References

[1]Hohmann, W., *Die Erreichbarkeit der Himmelskörper,* Oldenbourg, Munich, 1925; also, NASA Technical Translation F-44, 1960.

[2]Lawden, D. F., *Optimal Trajectories for Space Navigation,* Butterworths, London, 1963, Chap. 6.

[3]Barrar, R. B., "An Analytic Proof that the Hohmann-Type Transfer is the True Minimum Two-Impulse Transfer," *Astronautica Acta,* Vol. 9, No. 1, 1963, pp. 1–11.

[4]Hazelrigg, G. A., Jr., "The Use of Green's Theorem to Find Globally Optimal Solutions to a Class of Impulsive Transfers," American Astronomical Society, AAS Paper 68-092, Sept. 1968.

[5]Marec, J. P., *Optimal Space Trajectories,* Elsevier, Amsterdam, 1979, Chap. 2, pp. 21–27.

[6]Battin, R. H., *An Introduction to the Mathematics and Methods of Astrodynamics,* AIAA Education Series, AIAA, New York, 1987, pp. 529–530.

[7]Palmore, J. I., "An Elementary Proof of the Optimality of Hohmann Transfers," *Journal of Guidance, Control, and Dynamics,* Vol. 7, No. 5, 1984, pp. 629–630.

[8]Edelbaum, T. N., "How Many Impulses?" *Astronautics and Aeronautics,* Nov. 1967, pp. 64–69.

Symptom of Payload-Induced Flight Instability

Charles H. Murphy*
U.S. Army Ballistic Research Laboratory, Aberdeen Proving Ground, Maryland 21005

Introduction

THE angular motion of a projectile in flight can usually be represented as the sum of two coning motions. The stabil-

Received May 16, 1991; revision received Sept. 26, 1991; accepted for publication Sept. 27, 1991.

*Chief, Launch and Flight Division, SLCBR-LF. Fellow AIAA.