JOURNAL OF GUIDANCE, CONTROL, AND DYNAMICS
Vol. 21, No. 1, January–February 1998

Optimal, Low-Thrust, Earth–Moon Orbit Transfer

Albert L. Herman*
Spectrum Astro, Inc., Gilbert, Arizona 85233
and
Bruce A. Conway†
University of Illinois at Urbana–Champaign, Urbana, Illinois 61801

Optimal, low-thrust, Earth–moon orbit transfers are found using the method of collocation with nonlinear programming. The initial spacecraft Earth orbit is arbitrary; the final lunar orbit is arbitrary. The moon is in its actual orbit. The vehicle dynamics include the effects of the moon's gravity during the Earth-departure phase and the effects of the Earth's gravity during the lunar-arrival phase. Total transfer time is minimized. However, because the propulsion system operates continuously, i.e., no coast arc is allowed, the trajectory is also propellant minimizing. A very low initial thrust acceleration of 10^{-4} *g* yields flight times of approximately 32 days requiring many revolutions of both the Earth and the moon. Ignoring third-body gravity, i.e., solving the problem as two coupled two-body problems, changes the optimal trajectory only slightly, for example, decreasing the travel time by a few hours. The optimal trajectory is also insensitive to change in engine specific impulse so long as the same initial thrust acceleration magnitude is used.

Introduction

THE low thrust produced by an advanced propulsion system requires a vehicle using such a system to gradually change its orbit radius over several or many revolutions. Thus, the transfer does not resemble the elliptically shaped trajectory of the high-thrust or impulsive-thrust case. Instead, the resulting trajectory takes on a spiral shape where the orbit radius slowly increases (or decreases) as the spacecraft revolves about the attracting body. Such spiral orbit changes may take months to complete. Therefore, and because the propellant fraction is small, the minimum-time orbit transfer is often sought.

Low-thrust trajectory optimization problems are difficult to solve in part because the total transfer time is long. Computing optimal, low-thrust trajectories is difficult due to the sensitivity of a trajectory solution to boundary conditions. Typical methods for numerically solving optimal control problems work well with problems where the total time is relatively short. In this work, a new method for solving problems with relatively long total times, such as low-thrust orbit transfers, is used to solve for optimal, low-thrust, Earth–moon transfers.

Relative to the volume of analysis of Earth–moon orbit transfers using conventional propulsion systems, the number of studies of Earth–moon transfers using advanced propulsion is small. Aston[1] demonstrated the merits and feasibility of using low-thrust propulsion to ferry cargo between low Earth orbit (LEO) and low lunar orbit (LLO). He makes clear that the development of the optimal trajectories for such missions requires sophisticated system and mission studies. One early study was performed by Stuhlinger.[2] In his work, the low-thrust trajectory consists of separate Earth and moon two-body trajectory segments patched together at the sphere of influence of the moon. More recent work on optimal, low-thrust, Earth–moon orbit transfers has been performed by Enright and Conway.[3] Their study used a two-body, inverse-square gravitational field model for spacecraft motion about a central body. They solved minimum-fuel, low-thrust, Earth–moon transfer problems. Golan and Breakwell[4] investigated minimum-fuel, fixed flight time, Earth–moon transfers for power-limited spacecraft. The trajectory is found by matching an Earth spiral to a moon spiral at some intermediate distance.

Pierson and Kluever[5] find minimum-fuel, planar transfers from a circular LEO to a circular LLO. The Earth–moon trajectory is governed by the restricted three-body problem dynamics. The optimal Earth–moon transfer problem is solved by formulating and successively solving a hierarchy of subproblems, resulting in a three-stage solution process. Kluever and Pierson[6] improved the solution method in a second paper, doing away with the three-stage solution process, and solved minimum-fuel, three-dimensional trajectories from a circular LEO (assumed to lie in the lunar orbit plane) to an inclined circular LLO. The minimum-fuel transfer to a polar lunar orbit is obtained by successively solving a sequence of fixed lunar orbit inclination problems. In both papers, a relatively large thrust acceleration is used so that a translunar coast is fuel optimizing.[5,6]

In this work, optimal, low-thrust, Earth–moon orbit transfers are generated for the case where the initial spacecraft Earth orbit is arbitrary and the moon is in its actual orbit. The vehicle dynamics include the effects of the moon's gravity during the Earth-departure phase and the effects of the Earth's gravity during the lunar-arrival phase. Thus, this problem is a restricted three-body problem. Because the propulsion system requires only a small fraction of the total spacecraft mass and because the total transfer time is relatively long (on the order of 30 days), the total transfer time is minimized rather than minimizing the propellant usage during the transfer. However, because the propulsion system operates continuously, i.e., no coast arc is allowed, the trajectory is also propellant minimizing. The optimizer may choose the time of departure from circular orbit about the Earth. This variable would not influence the trajectory if the moon was assumed to be in a circular orbit but does influence the trajectory when the moon is assumed to have its true eccentric orbit because it determines the Earth–moon distance at the time of departure. The system model used in this work differs significantly from that used in Refs. 5 and 6 or by Golan and Breakwell.[4] Here the initial orbit of the spacecraft is arbitrary, the moon is in an eccentric and inclined orbit, and the thrust acceleration is significantly smaller, 10^{-4} *g*, 30 times smaller than that assumed by Refs. 5 and 6 and 10 times smaller than that used by Ref. 4. Most significantly, using collocation with nonlinear programming (NLP) allows the problem to be solved in one step for arbitrary final lunar orbit inclinations, i.e., the problem does not need to be divided into stages that require matching at one or more points.

The smaller thrust acceleration magnitude makes this problem more difficult, in principle, because the size of the problem, after it has been discretized and converted to an NLP problem, depends directly on the number of revolutions of the two bodies that are required. The optimal trajectories in this work require about 9 revolutions of the Earth for escape from geosynchronous orbit (GEO)

Received Feb. 3, 1997; revision received June 20, 1997; accepted for publication June 25, 1997.

*Mission Analyst, Advanced Space Systems. E-mail: albert.herman@specastro.com. Senior Member AIAA.

†Professor, Department of Aeronautical and Astronautical Engineering. E-mail: bconway@uiuc.edu. Associate Fellow AIAA.

and 17 revolutions for capture into and circularization of the lunar orbit. References 5 and 6, using a thrust acceleration 30 times larger, need 12 revolutions of the Earth for escape of their spacecraft from LEO but only 2 revolutions on arrival into LLO.

This paper begins by defining the system model used in solving this problem. This model is written using equinoctial elements as state variables,[7,8] yielding a nonsingular formulation even for cases of circular or equatorial orbits. A seventh-degree Gauss–Lobatto system constraint formulation[9] is used in the discretization of the continuous optimal control problem.

Method

The Earth–moon transfer problem solved involves two phases. In the first phase the propulsion system gradually increases the spacecraft's orbital energy relative to the Earth, resulting in a multirevolution orbit where the distance between the Earth and spacecraft increases with time. During the second phase of the transfer, the propulsion system gradually decreases the spacecraft's orbital energy relative to the moon. Thus, the arrival phase trajectory also takes on the form of a spiral. The governing equations are modeled relative to an Earth-centered inertial (ECI) system during the departure phase and relative to a moon-centered nonrotating (MCNR) system during the arrival phase. The state of the spacecraft at the end of the departure phase, i.e., the absolute position, velocity, and mass of the spacecraft, is the same as the state of the spacecraft at the beginning of the arrival phase. Thus, the equinoctial element values at the end of the departure phase and beginning of the arrival phase are directly related through a transformation of variables from the ECI system to the MCNR system, resulting in discontinuities in the state variables. This transformation may be rewritten as a constraint, resulting in a set of boundary conditions relating the states at the end of the departure phase to those at the beginning of the arrival phase. In addition, the governing equations are written in terms of localized units known as canonical units, for which the gravitational constant $\mu = 1$ in both the geocentric and lunicentric systems; therefore, the equations of motion are discontinuous as well. The motion of the moon is known as a function of time, resulting in the governing equations being an explicit function of time. Finally, the perturbing acceleration of the moon during the Earth-departure phase and the perturbing acceleration of the Earth during the lunar-arrival phase are modeled.

To avoid singularities that result from using the classical elliptic elements for circular or equatorial orbits, an alternative set of parameters is used. These parameters are the equinoctial elements, derived by Broucke and Cefola[7] and described by Battin.[8] With the equinoctial elements, the mean longitude ℓ and the parameters P_1, P_2, Q_1, and Q_2 replace the classical orbit elements e, i, Ω, ω, and f. The parameters P_1 and P_2 are defined by

$$P_1 = e \sin \varpi \tag{1}$$

and

$$P_2 = e \cos \varpi \tag{2}$$

where ϖ is the longitude of pericenter defined by $\varpi = \Omega + \omega$. The parameters Q_1 and Q_2 are defined by

$$Q_1 = \tan\left(\tfrac{1}{2}i\right) \sin \Omega \tag{3}$$

and

$$Q_2 = \tan\left(\tfrac{1}{2}i\right) \cos \Omega \tag{4}$$

The mean longitude ℓ is defined by $\ell = \varpi + M = \varpi + n\Delta t$, where M is the mean anomaly defined by $M = E - e \sin E$ (which is Kepler's equation) and E is the eccentric anomaly. The parameter n is the mean orbital motion and Δt is the time elapsed since periapse passage. Now, the eccentric longitude K is defined by $K = \varpi + E$. Substituting for $M = \ell - \varpi$ and $E = K - \varpi$, Kepler's equation becomes

$$\ell = K + P_1 \cos K - P_2 \sin K \tag{5}$$

which is the augmented form of Kepler's equation.[7,8] The true longitude L is defined as $L = \varpi + f$, where f is the true anomaly and L is related to the eccentric longitude K through

$$\sin L = (a/r)\Big(\big\{1 - [a/(a+b)]P_2^2\big\} \sin K + [a/(a+b)]P_1 P_2 \cos K - P_1\Big) \tag{6}$$

and

$$\cos L = (a/r)\Big(\big\{1 - [a/(a+b)]P_1^2\big\} \cos K + [a/(a+b)]P_1 P_2 \sin K - P_2\Big) \tag{7}$$

where

$$\frac{a}{r} = \frac{1}{1 - P_1 \sin K - P_2 \cos K}$$

and

$$\frac{a}{a+b} = \frac{1}{1 + \sqrt{1 - P_1^2 - P_2^2}}$$

The classical orbit elements e, i, Ω, and ω may be recovered from the relationships

$$e^2 = P_1^2 + P_2^2 \tag{8}$$

$$\tan^2\left(\tfrac{1}{2}i\right) = Q_1^2 + Q_2^2 \tag{9}$$

$$\tan \varpi = P_1/P_2 \tag{10}$$

and

$$\tan \Omega = Q_1/Q_2 \tag{11}$$

with $\omega = \varpi - \Omega$. True anomaly f may be recovered by solving Eq. (5) for the eccentric longitude K given ℓ, P_1, and P_2 and then using Eqs. (6) and (7) to compute a value for the true longitude L. Then, $f = L - \varpi$. Solving the augmented form of Kepler's equation (5) for the eccentric longitude K proved to be more difficult than solving Kepler's equation for the eccentric anomaly E. Therefore, a specialized algorithm was developed, which incorporates a nonderivative, dissection search technique based on the Fibonacci sequence with a faster derivative-based search technique.[9] The dissection technique is used to produce bounds on the range of K within which a solution to Eq. (5) lies so that the derivative search technique will quickly converge. This two-step algorithm has proven to be very robust in solving Eq. (5), the augmented form of Kepler's equation.

Equations of Motion

The equations of motion for the time rates of change of the equinoctial elements may be found in Refs. 7 and 8 and are given next. The time rates of change of the semimajor axis a, mean longitude ℓ, and parameters P_1, P_2, Q_1, and Q_2 are

$$\frac{\mathrm{d}a}{\mathrm{d}t} = 2\frac{a^2}{h}\left[(P_2 \sin L - P_1 \cos L)A_R + \frac{p}{r}A_T\right] \tag{12}$$

$$\frac{\mathrm{d}\ell}{\mathrm{d}t} = n - \frac{r}{h}\left\{\left[\frac{a}{a+b}\left(\frac{p}{r}\right)(P_1 \sin L + P_2 \cos L) + 2\frac{b}{a}\right]A_R + \frac{a}{a+b}\left(1 + \frac{p}{r}\right)(P_1 \cos L - P_2 \sin L)A_T + (Q_1 \cos L - Q_2 \sin L)A_N\right\} \tag{13}$$

$$\frac{\mathrm{d}P_1}{\mathrm{d}t} = \frac{r}{h}\left\{-\frac{p}{r}\cos L A_R + \left[P_1 + \left(1 + \frac{p}{r}\right)\sin L\right]A_T - P_2(Q_1 \cos L - Q_2 \sin L)A_N\right\} \tag{14}$$

$$\frac{dP_2}{dt} = \frac{r}{h}\left\{\frac{p}{r}\sin L A_R + \left[P_2 + \left(1 + \frac{p}{r}\right)\cos L\right]A_T + P_1(Q_1 \cos L - Q_2 \sin L)A_N\right\} \tag{15}$$

$$\frac{dQ_1}{dt} = \frac{r}{2h}\left(1 + Q_1^2 + Q_2^2\right)\sin L A_N \tag{16}$$

and

$$\frac{dQ_2}{dt} = \frac{r}{2h}\left(1 + Q_1^2 + Q_2^2\right)\cos L A_N \tag{17}$$

where $b = a\sqrt{(1 - P_1^2 - P_2^2)}$, which is the semiminor axis, and $h = nab$, which is the specific angular momentum. The useful ratios p/r and r/h are defined by $p/r = 1 + P_1 \sin L + P_2 \cos L$ and

$$\frac{r}{h} = \frac{h}{\mu(1 + P_1 \sin L + P_2 \cos L)}$$

respectively. The mean orbital motion is given by $n = \sqrt{(\mu/a^3)}$. The variables A_R, A_T, and A_N represent the perturbing accelerations, defined in the next section.

Perturbing Accelerations

If the perturbing accelerations A_R, A_T, and A_N are all identically equal to zero, then the differential equations (12–17) describe the trajectory of a spacecraft traveling in an orbit whose elements a, P_1, P_2, Q_1, and Q_2 are fixed, with mean longitude ℓ advancing at a constant rate. However, in this work, a spacecraft with a low-thrust propulsion system is modeled. Thus, the perturbing accelerations are small but significant, causing the orbit to change over time. In addition, the gravitational attraction of the moon during the Earth-departure phase of a transfer and the gravitational attraction of the Earth during the lunar-arrival phase are modeled as perturbations. Thus,

$$A_R = \Gamma \cos\alpha \sin\beta + \Im_R \tag{18}$$

$$A_T = \Gamma \cos\alpha \cos\beta + \Im_T \tag{19}$$

$$A_N = \Gamma \sin\alpha + \Im_N \tag{20}$$

where Γ is the spacecraft's instantaneous thrust acceleration magnitude, α is the thrust acceleration pointing angle measured out of the instantaneous orbit plane, and β is the in-plane thrust acceleration pointing angle. The angle α is positive if the thrust has a component in the direction of the orbit angular momentum vector. The angle β is zero if the thrust is directed normal to the radius vector and is positive if the thrust has a radially outward component. The thrust pointing angles α and β are the control variables for this problem. The variables $\Im_R$, $\Im_T$, and $\Im_N$ are the third-body perturbing accelerations in the radial, transverse, and normal directions relative to the instantaneous orbit plane.

The change in thrust acceleration as propellant is consumed is modeled by

$$\frac{d\Gamma}{dt} = \frac{1}{c}\Gamma^2 \tag{21}$$

where c is the effective exhaust velocity of the propulsion system. The effective exhaust velocity is defined by $c = g_0 I_{sp}$, where g_0 is the gravity magnitude at the surface of the Earth, which is approximately 9.81 m/s^2, and I_{sp} is the specific impulse of the propulsion system. Thus, different low-thrust propulsion systems with varying specific impulses may be analyzed.

Defining $\Im = \{\Im_R, \Im_T, \Im_N\}^T$, then,

$$\Im_M = -\mu_M\left[\frac{\boldsymbol{r}_{MS}}{r_{MS}^3} + \frac{\boldsymbol{r}_{EM}}{r_{EM}^3}\right] \tag{22}$$

is the perturbing acceleration of the moon during the Earth-departure phase and

$$\Im_E = -\mu_E\left[\frac{\boldsymbol{r}_{ES}}{r_{ES}^3} - \frac{\boldsymbol{r}_{EM}}{r_{EM}^3}\right] \tag{23}$$

is the perturbing acceleration of the Earth during the lunar-arrival phase.[10] Here, $\boldsymbol{r}_{MS}$ and $\boldsymbol{r}_{ES}$ are the positions of the spacecraft relative to the moon and Earth, respectively, and $\boldsymbol{r}_{EM}$ is the position of the moon relative to the Earth.

The position of a spacecraft relative to the local coordinate system may be computed using

$$\boldsymbol{r} = [\boldsymbol{R}]\begin{bmatrix} r\cos L \\ r\sin L \\ 0 \end{bmatrix} \tag{24}$$

where

$$[\boldsymbol{R}] = \frac{1}{1 + Q_1^2 + Q_2^2} \times \begin{bmatrix} 1 - Q_1^2 + Q_2^2 & 2Q_1Q_2 & 2Q_1 \\ 2Q_1Q_2 & 1 + Q_1^2 - Q_2^2 & -2Q_2 \\ -2Q_1 & 2Q_2 & 1 - Q_1^2 - Q_2^2 \end{bmatrix} \tag{25}$$

and $r = a(1 - P_1 \sin K - P_2 \cos K)$. Thus, Eqs. (24) and (25) may be used to find either $\boldsymbol{r}_{MS}$ or $\boldsymbol{r}_{ES}$. If used to find $\boldsymbol{r}_{ES}$, equinoctial elements appropriate to an Earth-centered orbit are used in Eqs. (24) and (25) and $\boldsymbol{r}_{ES}$ will then be expressed on an ECI basis. If used to find $\boldsymbol{r}_{MS}$, equinoctial elements appropriate to a moon-centered orbit are used in Eqs. (24) and (25) and $\boldsymbol{r}_{MS}$ will then be expressed on an MCNR basis, which is a basis that translates but does not rotate as the moon travels about the Earth. The position of the moon relative to the Earth, $\boldsymbol{r}_{EM}$, is computed in the ECI reference frame using

$$\boldsymbol{r}_{EM} = [\boldsymbol{S}]\begin{bmatrix} a_M(\cos E_M - e_M) \\ a_M\sqrt{1 - e_M^2}\sin E_M \\ 0 \end{bmatrix} \tag{26}$$

where

$$[\boldsymbol{S}] = \begin{bmatrix} c\Omega_M c\omega_M - s\Omega_M s\omega_M s\iota_M & -c\Omega_M s\omega_M - s\Omega_M c\omega_M c\iota_M & s\Omega_M s\iota_M \\ s\Omega_M c\omega_M + c\Omega_M s\omega_M c\iota_M & -s\Omega_M s\omega_M + c\Omega_M c\omega_M c\iota_M & -c\Omega_M s\iota_M \\ s\omega_M s\iota_M & c\omega_M s\iota_M & c\iota_M \end{bmatrix} \tag{27}$$

as described in Ref. 10. In Eq. (27), c denotes cos( ) and s denotes sin( ). Here, the lunar orbital elements a_M and e_M are assumed to be $60.27R_\oplus$ and 0.0549, respectively, the actual mean values, where $R_\oplus$ is the mean equatorial radius of the Earth. The lunar orbit inclination with respect to the equatorial plane varies between approximately 18.2 and 28.5 deg; an inclination of 25 deg is chosen for the example transfers presented here. The positions of the node and the line of apsides circulate; for this model $\Omega_M = \omega_M = 0$ is assumed for convenience; the choice does not affect the difficulty of solving the problem. Of course, the exact values for all of the orbit elements for a given departure date would be used if that date were known. In any event, the initial true longitude of the moon does not affect the transfer because the spacecraft is allowed to coast until the relative position of the spacecraft and moon is such as to minimize the powered flight time. The eccentric anomaly of the lunar orbit, E_M, is determined from Kepler's equation:

$$M_M = E_M - e_M \sin E_M \tag{28}$$

where $M_M = n_M t$ and is known as the lunar mean anomaly, where $n_M = \sqrt{(\mu_E / a_M^3)}$ is the lunar mean orbital motion. In this work, the value of the lunar mean anomaly at zero time is considered to be zero. Equation (28) is solved using the Laguerre method of Conway.[11]

With $\boldsymbol{r}_{\mathrm{ES}}$ and $\boldsymbol{r}_{\mathrm{EM}}$ found, as described, on the ECI basis, $\mathfrak{I}_M$ found as the right-hand side of Eq. (22) will also be expressed on the ECI basis. The gravitational perturbation $\mathfrak{I}_M$ must then be resolved into components in the spacecraft-fixed radial, tangential, and normal directions for use in Eqs. (18–20). This is accomplished by the transformation

$$\mathfrak{I}_M \text{ (local)} = [\boldsymbol{T}]^T [\boldsymbol{R}]^T \mathfrak{I}_M \text{ (ECI)} \tag{29}$$

where

$$[\boldsymbol{T}] = \begin{bmatrix} \cos L & -\sin L & 0 \\ \sin L & \cos L & 0 \\ 0 & 0 & 0 \end{bmatrix} \tag{30}$$

Similarly, when $\boldsymbol{r}_{\mathrm{MS}}$ is found as described, on the MCNR basis, and with $\boldsymbol{r}_{\mathrm{EM}}$ determined on the MCNR basis through

$$\boldsymbol{r}_{\mathrm{EM}} = \begin{bmatrix} a_M(\cos E_M - e_M) \\ a_M\sqrt{1-e_M^2}\sin E_M \\ 0 \end{bmatrix} \tag{26a}$$

then $\mathfrak{I}_E$ found as the right-hand side of Eq. (23) will also be expressed on the MCNR basis. As described in the preceding paragraph, the gravitational perturbation $\mathfrak{I}_E$ must then be resolved into components in the spacecraft-fixed radial, transverse, and normal directions for use in Eqs. (18–20). This is accomplished again by the transformations (29) and (30), with the only difference being that the true longitude L is now determined using Eqs. (6) and (7) using equinoctial variables appropriate to the moon-centered orbit.

Earth-Phase to Lunar-Phase Boundary Conditions

As mentioned, the system state variables are the equinoctial variables plus the thrust acceleration, i.e., $[a, P_1, P_2, Q_1, Q_2, \ell, \Gamma]$. At the point where the trajectory switches from the Earth-departure phase to the lunar-arrival phase, which is chosen to be the first point in the discretized time history such that the spacecraft enters within 25 lunar radii of the moon, a transformation is made from equinoctial elements appropriate to Earth-centered motion (perturbed by the moon) to equinoctial elements appropriate to moon-centered motion (perturbed by the Earth). Note that, in this work, unlike models that use a patched-conic approximation, this switch may be made at any point without affecting the dynamics because a three-body gravitational model is always used. One method for accomplishing the transformation uses the continuity of the absolute position and velocity of the spacecraft. At any time,

$$\boldsymbol{r}_{\mathrm{ES}} = \boldsymbol{r}_{\mathrm{EM}} + \boldsymbol{r}_{\mathrm{MS}} \tag{31}$$

and

$$\boldsymbol{v}_{\mathrm{ES}} = \boldsymbol{v}_{\mathrm{EM}} + \boldsymbol{v}_{\mathrm{MS}} \tag{32}$$

where $\boldsymbol{v}_{\mathrm{ES}}$ is the velocity vector of a spacecraft relative to an ECI coordinate system, $\boldsymbol{v}_{\mathrm{EM}}$ is the velocity vector of the moon relative to the ECI system, and $\boldsymbol{v}_{\mathrm{MS}}$ is the velocity vector of the spacecraft relative to an MCNR coordinate system. Vectors $\boldsymbol{r}_{\mathrm{ES}}$ and $\boldsymbol{r}_{\mathrm{MS}}$ may be found from Eqs. (24) and (25) but with each position vector found as a function of the local equinoctial elements. The velocity vectors $\boldsymbol{v}_{\mathrm{ES}}$ and $\boldsymbol{v}_{\mathrm{MS}}$ of the spacecraft may be determined in terms of the local equinoctial elements using

$$\boldsymbol{v} = [\boldsymbol{R}]\left(\frac{\mu}{h}\right)\begin{bmatrix} -P_1 - \sin L \\ P_2 + \cos L \\ 0 \end{bmatrix} \tag{33}$$

The lunar position $\boldsymbol{r}_{\mathrm{EM}}$ may be computed using either Eq. (26) or Eq. (26a), and the lunar velocity $\boldsymbol{v}_{\mathrm{EM}}$ may be determined using

$$\boldsymbol{v}_{\mathrm{EM}} = [\boldsymbol{S}]\begin{bmatrix} \dfrac{-\mu_E}{h_M}\sqrt{1-e_M^2}\dfrac{\sin E_M}{1-e_M\cos E_M} \\ \dfrac{\mu_E}{h_M}\left(1-e_M^2\right)\dfrac{\cos E_M}{1-e_M\cos E_M} \\ 0 \end{bmatrix} \tag{34}$$

where h_M is the specific angular momentum of the moon. Using the transformation matrix in Eq. (27), the vectors $\boldsymbol{r}_{\mathrm{MS}}$ and $\boldsymbol{v}_{\mathrm{MS}}$, which are originally expressed on the MCNR basis, may be expressed on the ECI basis. Then Eqs. (31) and (32) yield six nonlinear constraint equations involving the two sets of equinoctial variables; enforcing these constraints when the problem is solved accomplishes the transformation of variables from the Earth-departure phase to the moon-arrival phase. Of course, the thrust acceleration Γ is continuous across this transformation.

However, the thrust acceleration magnitude must be computed in terms of the local coordinates. Because canonical units ($\mu = 1$) are used throughout, the thrust acceleration magnitude conversion becomes one of unit conversion. The length units (LU) and time units (TU) in each phase are chosen such that the gravitational constant μ is unity. (In these canonical units, 2π TUs would represent the period of a circular orbit at the surface of the attracting body.) The LU and TU for the Earth are 1 $\mathrm{LU}_\oplus = 6378$ km and 1 $\mathrm{TU}_\oplus = 806.8$ s, respectively, and for the moon the LU and TU are 1 $\mathrm{LU}_M = 1738$ km and 1 $\mathrm{TU}_M = 1035$ s, respectively. Thus, the thrust acceleration magnitude at the end of the departure phase and the beginning of the arrival phase are related by

$$\Gamma_\oplus\left(\mathrm{LU}_\oplus/\mathrm{TU}_\oplus^2\right) = \Gamma_M\left(\mathrm{LU}_M/\mathrm{TU}_M^2\right) \tag{35}$$

To correctly account for the differences between units used for the quantities in the constraint equations (31) and (32), the units for distance and time for the lunar-arrival phase were converted to the units used for the Earth-departure phase.

Initial and Final Orbit Conditions

The initial orbit of the spacecraft is assumed to be a circular, GEO inclined at an angle of 28.5 deg to Earth's equatorial plane and having a radius of $6.63R_\oplus$. Thus, initially, $a = 6.63$, $e = 0$, and $i = 28.5$ deg with true anomaly f free to be optimally chosen. Because the initial orbit is inclined to the Earth equatorial plane, it is three dimensional irrespective of the initial values of the longitude of the ascending node and argument of perigee. Thus, the initial values of the longitude of the ascending node and argument of pericenter are assumed to be zero, i.e., $\Omega = \omega = 0$. The initial magnitude of the thrust acceleration is arbitrarily chosen to be 10^{-4} LU/TU^2 ($= 10^{-4}$ g). Therefore, using Eqs. (1–4), the initial conditions for the states in the Earth-departure phase of the transfer are determined to be

$$\begin{aligned} \boldsymbol{x}_I &= \{a, \ell, P_1, P_2, Q_1, Q_2, \Gamma\}^T \\ &= \{6.63, \text{free}, 0, 0, 0, 0.254, 10^{-4}\}^T \end{aligned}$$

where the initial value of mean longitude ℓ in the departure phase of the transfer is free to be optimally chosen due to the free initial value of f. The final orbit of the spacecraft is chosen to be a circular lunar polar orbit with a specified final orbit radius, for example, a final radius of $2R_M$, where R_M is the moon's mean equatorial radius. Therefore, the final conditions for the states in the lunar-arrival phase are

$$\begin{aligned} \boldsymbol{x}_F &= \{a, \ell, P_1, P_2, Q_1, Q_2, \Gamma\}^T \\ &= \{\text{specified}, \text{free}, 0, 0, 0, 1, \text{free}\}^T \end{aligned}$$

where, again, the final value of ℓ in the arrival phase of the transfer is free to be optimally chosen due to the free final value of f. The final value of Γ is also a free variable. However, this final value must be appropriate considering an initial value of the thrust acceleration magnitude of 10^{-4} and the amount of propellant consumed during the (optimized) transfer time.

Direct Method

There are many possible ways of solving optimal control problems using direct methods. All such solution methods discretize or parametrize the time histories of the state or control variables (or

both) and then construct equations of constraint that enforce the system differential equations, at various interior points, and any specified boundary conditions. Perhaps the most familiar direct method is that of Hargraves and Paris[12] because it is the basis of the OTIS program.[13] In this method the constraint equation or defect (in their terms[12]) causes the system differential equations to be integrated across each time segment according to Simpson's rule.[3] Formulating these equations of constraint using higher-order integration rules has the same advantages as using higher-order rules, e.g., fourth-order Runge–Kutta, for numerical integration; fewer time steps need be used to achieve the same accuracy. This has been demonstrated in numerical experiments by Herman[9] and Herman and Conway.[14] For direct methods, such as the one used here, in which the discretized optimization problem is solved as an NLP problem, the reduction in the number of time steps yields a direct reduction in the number of NLP variables.

The solution method used in this work is a seventh-degree Gauss–Lobatto collocation method described by Herman and Conway[14] and Herman.[9] In this method an approximate solution to a system of first-order differential equations of the form $\dot{x} = f(x, t)$ is constructed on a time history that has been discretized. Parameters are used to represent values of the states x_i at discrete times t_i. Collocation constraints are then formulated by first constructing seventh-degree polynomials using the state parameter values and the system equations evaluated using the state parameters and discrete values of the control variables. Requiring the first derivative of the seventh-degree polynomial representing a given state to equal the system differential equation for that state at three points within each time segment yields three constraint equations. The locations of the three collocation points are chosen to minimize the error of the implicit integration. Note that this process of creating constraints is qualitatively similar to the direct method of Hargraves and Paris[12] but that method uses only a third-degree polynomial and yields just one constraint equation per segment.

The discrete values of the state and control variables, the time at which powered flight is begun, and the final time become parameters for the optimizer to choose to minimize the objective function, which is the flight time. The constraint equations described, which enforce satisfaction of the system equations (12–17) and (21), are nonlinear in these parameters, as are the earlier described transformation equations (31) and (32) of the equinoctial orbit elements from an Earth-centered reference to a moon-centered reference as the spacecraft approaches the moon. Thus, the problem is an NLP problem, which we solve using the routine NZOPT.[15]

Initial values for the parameters used to discretize the solution time histories must be provided before the NLP algorithm can compute the first step in which to begin the optimization search process. In this research, the initial values of the states are computed using a trajectory that is integrated without concern for the boundary conditions or for optimality. The initialization trajectory for the departure phase is computed using starting conditions that correspond to a circular, geostationary orbit. In the arrival phase, the initialization trajectory is determined with a backward integration from a circular, equatorial orbit with a radius equal to the specified final orbit radius. In both phases, the control variables α and β are assumed identically zero throughout the initial trajectory computation. Thus, the thrust pointing direction is always transverse to the instantaneous orbit radius vector. The initialization trajectories for the departure and arrival phases are integrated until the semimajor axis reaches values of $60R_{\oplus}$ and $25R_M$, respectively. These initialization trajectories are two-dimensional trajectories that do not satisfy the boundary conditions relating the state of a spacecraft at the end of the departure phase to the state of the spacecraft at the beginning of the arrival phase. In addition, third-body effects are not considered in the computation of the initialization trajectories. Thus, the initial values of the discretization parameters for the state and control variables do not correspond to a feasible solution, let alone an optimal solution.

The optimizer, NZOPT, also requires the partial derivatives of the system constraints with respect to each of the free parameters. These are computed numerically using the IMSL routine F2JAC.[16] This numerical routine computes the Jacobian of a vector function using forward differences.

Results

In this section the results of the Earth–moon orbit transfer problem already described are presented. In all cases the number of subintervals used for the Earth-departure and lunar-arrival phases are 20 and 10, respectively. In addition, the final lunar radius is arbitrarily chosen to be $2R_M$, the I_{sp} was set to (a realistic value of) 5000 s, and third-body effects are included for case A. For case B, the third-body effects are not modeled (for comparison to the results of case A). Case C is the same as case A with the exception of I_{sp}, which is set to 2500 s. Finally, case D is the same as case A except that the final lunar orbit radius is specified to be $3R_M$, again for comparison purposes. These cases are summarized in Table 1.

There is very little difference in flight time (and, hence, in final mass) for cases A and B, whose only difference is the inclusion of third-body gravity. For case C, where the engine is assumed much less efficient but that is otherwise the same as case A, the flight time is reduced. This is because the propellant is consumed more rapidly to produce the same thrust, so that the less massive spacecraft of case C has a higher thrust acceleration at all times than the spacecraft of cases A or B. The spacecraft of case D uses the least propellant because it operates approximately two fewer days than the spacecraft of the other cases to insert into a higher lunar orbit.

The trajectories of case A are described in Figs. 1–11. The departure phase of the optimal orbit transfer is shown in Fig. 1 and is described by Figs. 2–4; the arrival phase is shown in Fig. 5 and is described by Figs. 6–11. The semimajor axis and orbit radius for the departure and arrival phases are shown in Figs. 2 and 6, respectively. The semimajor axis increases monotonically as energy is added to the orbit during the departure; it decreases as energy is removed from the orbit during arrival. Time histories of the orbit eccentricity for the departure and arrival phases are shown in Figs. 3 and 7, respectively. The inclination time history of the departure phase is shown in Fig. 3. Change in inclination occurs primarily near the end of the departure phase, as expected because the spacecraft is moving more slowly when distant from the Earth and inclination change is less costly. The inclination changes little during the arrival phase, as shown in Fig. 11, because the vehicle is in nearly a polar orbit as it begins to orbit the moon. (This is also clear from Fig. 5, where only the X–Z projection, in lunicentric coordinates, is shown.) Thrust acceleration pointing angles for the departure and arrival phases are shown in Figs. 4 and 8, respectively. The initial orbit transfer time, when the spacecraft begins powered flight, and the final orbit transfer time,

Table 1 Summary of Earth–moon orbit transfer cases

Case	Three body	I_{sp}, s	Final lunar orbit radius, R_M	Ratio of final/initial mass, %	Time of flight, days
A	Yes	5000	2	94.2	33.54
B	No	5000	2	94.3	33.32
C	Yes	2500	2	88.7	32.67
D	Yes	5000	3	94.6	31.15

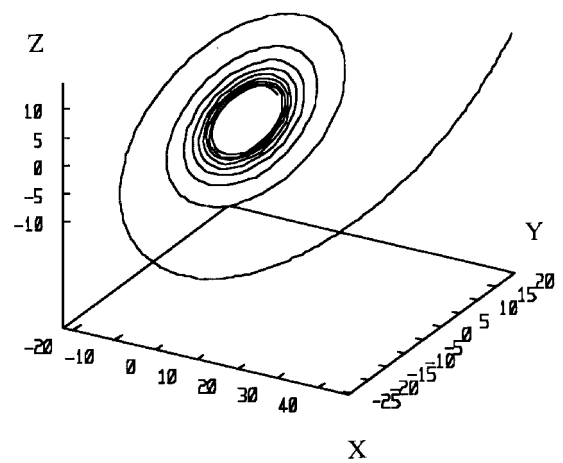

Fig. 1 Case A departure phase: three-dimensional view of the trajectory.

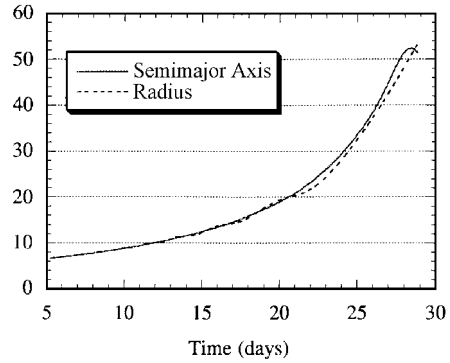

Fig. 2 Case A departure phase: semimajor axis and radius time histories (unit is Earth radius).

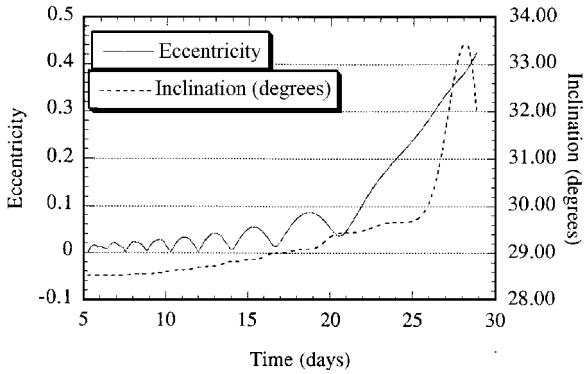

Fig. 3 Case A departure phase: eccentricity and inclination time histories.

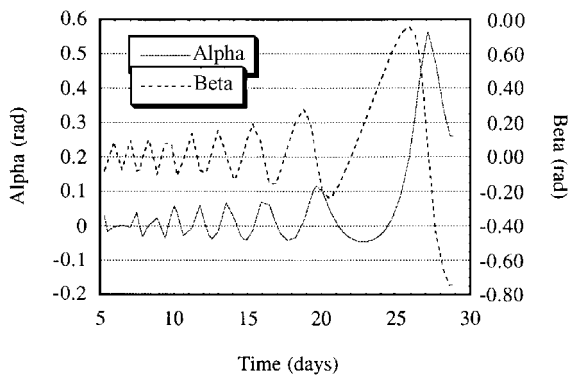

Fig. 4 Case A departure phase: time histories of thrust acceleration pointing angles; angle β is measured in-plane; α is measured normal to the instantaneous orbit plane.

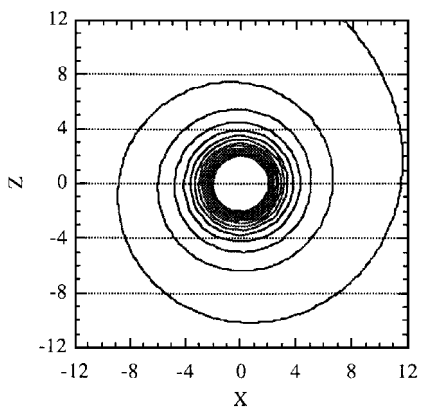

Fig. 5 Case A arrival phase: projection in X–Z lunicentric equatorial coordinates.

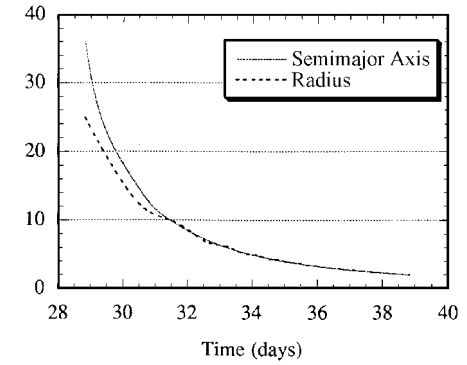

Fig. 6 Case A arrival phase: semimajor axis and radius time histories (unit is moon radius).

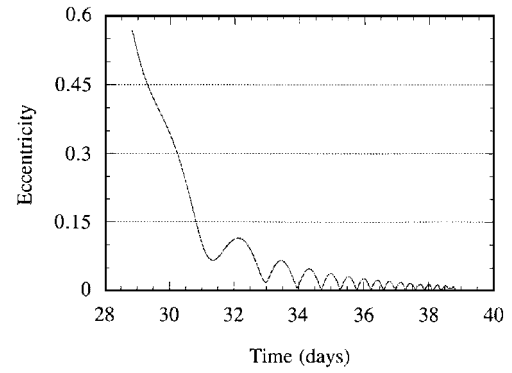

Fig. 7 Case A arrival phase: eccentricity time history.

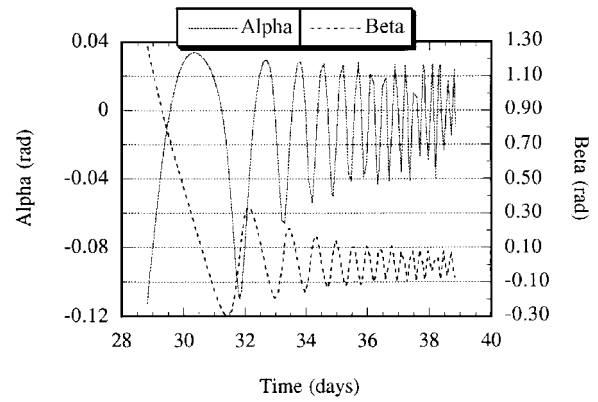

Fig. 8 Case A arrival phase: time histories of thrust acceleration pointing angles; angle β is measured in-plane; α is measured normal to the instantaneous orbit plane.

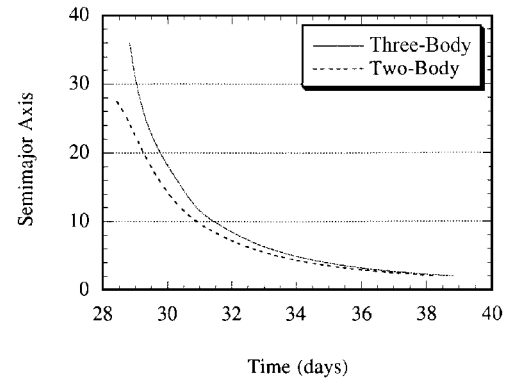

Fig. 9 Case A and B arrival phases: semimajor axis time history comparison (unit is moon radius).

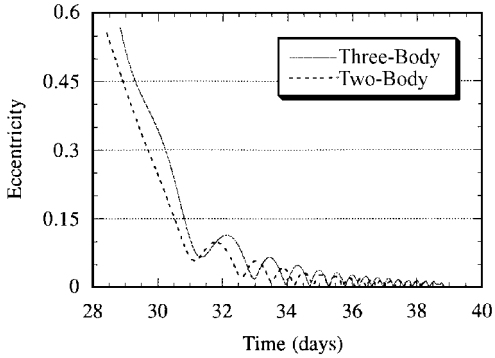

Fig. 10 Case A and B arrival phases: eccentricity time history comparison.

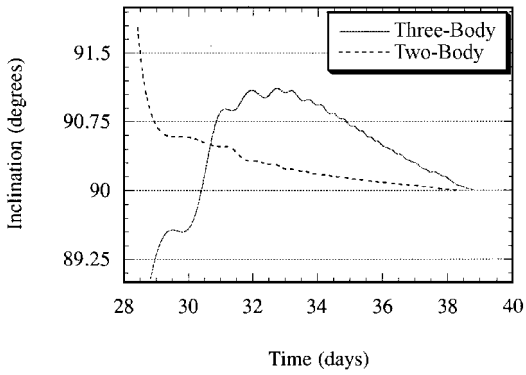

Fig. 11 Case A and B arrival phases: inclination time history comparison.

when the spacecraft completes the transfer, are free to be chosen by the NLP algorithm. In case A, the optimal orbit transfer is initiated at 5.286 days past the time of lunar perigee and continues until 38.83 days past perigee, resulting in a total transfer time of 33.55 days.

In case B, the third-body perturbing accelerations are not included in the system model, i.e., $\Im_M = \Im_E = \mathbf{0}$ in this case, which is otherwise the same as case A. Powered flight is initiated at 4.969 days past the lunar perigee and continues until 38.29 days past perigee, resulting in a total transfer time of 33.32 days, which is only 0.23 days less than that for the three-body case A. Cases A and B are compared in Figs. 9–11. The departure-phase semimajor axis time history for case B is almost indistinguishable from that of case A (shown in Fig. 2) and, thus, is not shown. This result was expected because the disturbing acceleration of the moon on the spacecraft is negligibly small while the spacecraft is near the Earth. The arrival-phase semimajor axis time history is shown in Fig. 9, in which the influence of the moon as the spacecraft approaches the moon is evident. The eccentricity and inclination time histories shown in Figs. 10 and 11 also show a significant difference when third-body effects are considered. However, the eccentricity time histories are essentially the same but shifted in time to reflect the fact that the optimal transfer, when third-body effects are included, begins approximately 0.3 day later than when third-body effects are ignored. In addition, the differences in orbit inclination in the arrival phase are exaggerated by the scaling of the figure. The results are much more completely presented in Ref. 9.

Conclusions

The method of collocation with NLP has successfully solved a trajectory optimization problem that is difficult due to the long flight time and the many-revolution departure and arrival trajectories. With this method the problem may be solved in one step for arbitrary final lunar orbit inclinations; the problem does not need to be divided into stages that require matching at one or more points, and that may yield a combined trajectory that is suboptimal. The Earth–moon orbit transfer problem solved is quite realistic. For instance, the spacecraft is considered to be initially in a plane that is inclined to the Earth equatorial plane and transfers to a final polar lunar orbit (the most useful orbit for lunar mapping). The orbit of the moon used in the model is the actual, eccentric orbit, in a plane inclined to the Earth's equatorial plane, resulting in a complex, three-dimensional problem. In addition, third-body effects are included during both the Earth-departure and lunar-arrival phases. Finally, representative electric propulsion specific impulses and thrust magnitudes are used, yielding total transfer times of approximately one month.

The perturbing effect of the Earth's gravity on the lunar-arrival trajectory is significant. The perturbing effect of the moon's gravity influences only the very end of the Earth-departure phase. A change in specific impulse does affect the orbit transfer with a significant change in transfer time as well as changes to the departure and arrival trajectories. A change in the final desired lunar orbit radius has no significant effect on the departure phase, and the only significant effect on the overall problem is a shorter transfer time.

The solution method applied successfully in this work should have no difficulty solving a problem with a larger initial low-thrust acceleration than the 10^{-4} g assumed here and, thus, involving fewer revolutions of the Earth and moon because the system equations of motion (implicitly) integrated in the determination of the optimal trajectory are the exact variational equations, i.e., no assumptions are made with regard to the magnitude of the variation of the orbit elements over each orbit. If the initial acceleration is instead made smaller, the method is likely, based on experience with its use, to solve the problem successfully. To maintain the same accuracy, it will be necessary to allocate more segments and, hence, more discrete states and controls; however, the seventh-degree Gauss–Lobatto method may provide sufficient accuracy for a problem with the same number of NLP variables and, hence, a coarser discretization.

Acknowledgment

We would like to thank the McDonnell–Douglas Astronautics Corporation for allowing us to use the NLP problem-solving code NZOPT.

References

[1]Aston, G., "Ferry to the Moon," *Aerospace America*, Vol. 25, No. 6, 1987, pp. 30–32.

[2]Stuhlinger, E., *Ion Propulsion for Spaceflight*, McGraw–Hill, New York, 1964.

[3]Enright, P. J., and Conway, B. A., "Discrete Approximations to Optimal Trajectories Using Direct Transcription and Nonlinear Programming," *Journal of Guidance, Control, and Dynamics*, Vol. 15, No. 4, 1992, pp. 994–1002.

[4]Golan, O. M., and Breakwell, J. V., "Minimum Fuel Lunar Trajectories for a Low-Thrust Power-Limited Spacecraft," AIAA Paper 90-2975, 1990.

[5]Pierson, B. L., and Kluever, C. A., "Three-Stage Approach to Optimal Low-Thrust Earth–Moon Trajectories," *Journal of Guidance, Control, and Dynamics*, Vol. 17, No. 6, 1994, pp. 1275–1282.

[6]Kluever, C. A., and Pierson, B. L., "Optimal Low-Thrust Three-Dimensional Earth–Moon Trajectories," *Journal of Guidance, Control, and Dynamics*, Vol. 18, No. 4, 1995, pp. 830–837.

[7]Broucke, R. A., and Cefola, P. J., "On the Equinoctial Orbit Elements," *Celestial Mechanics*, Vol. 5, No. 3, 1972, pp. 303–310.

[8]Battin, R. H., *An Introduction to the Mathematics and Methods of Astrodynamics*, AIAA Education Series, AIAA, New York, 1987, pp. 490–494.

[9]Herman, A. L., "Improved Collocation Methods with Application to Direct Trajectory Optimization," Ph.D. Thesis, Dept. of Aeronautical and Astronautical Engineering, Univ. of Illinois, Urbana–Champaign, IL, Sept. 1995.

[10]Prussing, J. E., and Conway, B. A., *Orbital Mechanics*, Oxford Univ. Press, New York, 1993, pp. 6–9.

[11]Conway, B. A., "An Improved Algorithm Due to Laguerre for the Solution of Kepler's Equation," *Celestial Mechanics*, Vol. 39, No. 2, 1986, pp. 199–211.

[12]Hargraves, C. R., and Paris, S. W., "Direct Trajectory Optimization Using Nonlinear Programming and Collocation," *Journal of Guidance, Control, and Dynamics*, Vol. 10, No. 4, 1987, pp. 338–342.

[13]Paris, S. W., Hargraves, C. R., and Martens, P. J., "Optimal Trajectories by Implicit Simulation (OTIS)," Version 2.0, WRDC-TR-90-3056, Boeing Aerospace Corp., Seattle, WA, Dec. 1991.

[14]Herman, A. L., and Conway, B. A., "Direct Optimization Using Collocation Based on Higher-Order Gauss–Lobatto Quadrature Rules," *Journal of Guidance, Control, and Dynamics*, Vol. 19, No. 3, 1996, pp. 592–599.

[15]Gill, P. E., Saunders, M. A., and Murray, W., "User's Guide for NZOPT 1.0: A FORTRAN Package for Nonlinear Programming," McDonnell–Douglas Aerospace Corp., Huntington Beach, CA, April 1993.

[16]IMSL Math/Library, "User's Manual," IMSL, Inc., Houston, TX, 1990.