JOURNAL OF GUIDANCE, CONTROL, AND DYNAMICS
Vol. 26, No. 2, March–April 2003

Closed-Loop Endoatmospheric Ascent Guidance

Ping Lu,* Hongsheng Sun,† and Bruce Tsai‡
Iowa State University, Ames, Iowa 50011-2271

A comprehensive treatment to the optimal atmospheric ascent problem of launch vehicles subject to path constraints and final condition constraints is presented. The development is particularly tailored for the purpose of the eventual realization of closed-loop endoatmospheric ascent guidance. It is demonstrated that the classical finite difference method for two-point boundary-value problems is well suited for solving the optimal ascent problem onboard. The performance and execution of the guidance algorithm are assessed with a series of open-loop and closed-loop tests using the vehicle data of a reusable launch vehicle. The test results render strong supporting evidence to the belief that closed-loop optimal endoatmospheric ascent guidance is now feasible.

I. Introduction

THE ascent guidance system of a rocket-powered launch vehicle determines the attitude commands and, when applicable, engine throttle command during the ascent of the vehicle. It is well known that whether or not the ascent trajectory is optimal can have a significant impact on propellant usage for a given payload, or on payload weight for the same gross vehicle weight. Consequently, ascent guidance commands are usually optimized in some fashion. In fact, ascent guidance is one of the most notable engineering fields where optimal control theory has found routine applications. Successful vacuum rocket guidance software based on the optimal control theory includes IGM for Saturn rockets[1] and PEG for the space shuttle.[2] These algorithms solve the optimal vacuum powered flight problem onboard in each guidance update cycle using the current condition as the initial condition of the solution. Therefore, the guidance strategy in effect is closed loop.

The open challenges of ascent guidance lie in the endoatmospheric portion of the flight. The presence of the aerodynamic forces, loads, and winds significantly complicates the optimal ascent problem, making the solution process much more difficult to converge reliably and sufficiently fast for onboard applications. For these reasons, typical current ascent guidance inside the atmosphere is open loop.[3] In such an approach, the guidance commands are generated offline, updated with the day-of-launch wind data before launch, and loaded into the launch vehicle for use during the ascent through the atmosphere. Whereas very successful in nominal ascent guidance, the open-loop approach inherently lacks the adaptive capability to handle contingencies and aborts, even with extensive offline planning at great costs. Open-loop guidance also does not possess the robustness necessary to cope with significant off-nominal conditions and system modeling uncertainty, especially for new launch vehicles for which little or no flight data are available. The required replanning and regeneration of the open-loop ascent guidance commands whenever any mission or system parameters change are costly in both developmental and operational phases of the launch vehicle.

A closed-loop ascent guidance algorithm could address all of the stated deficiencies of open-loop guidance. The search for a feasible algorithm to solve the optimal control problem onboard for closed-loop atmospheric ascent guidance dates back to the 1960s. The work by Brown and Johnson represents one of the earliest attempts in this direction.[4] In recent work that has stimulated renewed interest in this area, Leung and Calise,[5] Calise et al.,[6] and Gath and Calise[7] develop a hybrid approach to the problem. In this approach, the analytical solution of the optimal vacuum flight and numerical collocation for the atmospheric portion are combined. The vacuum solution serves as the initial guess for the atmospheric flight, and a homotopy method is used to phase in gradually the aerodynamic terms and path constraint-related terms. Their simulations demonstrate remarkable convergence speed and suggest that a closed-loop endoatmospheric ascent guidance algorithm is feasible and within reach.

Under the current NASA Space Launch Initiative program, closed-loop endoatmospheric ascent guidance is being examined for the second-generation reusable launch vehicle (RLV). In addition to better performance, robustness, and cost benefits, such a closed-loop ascent guidance system would potentially provide the capability for safe autonomous aborts. This paper represents another effort to bring closed-loop optimal endoatmospheric ascent guidance closer to reality. The main contributions of this paper are threefold:

1) The three-dimensional optimal ascent problem is comprehensively treated subject to the common path constraints and orbital insertion conditions, as are techniques to address a number of onboard implementation issues, some of which are unique to nonaxisymmetric launch vehicles (such as a lifting-entry RLV).

2) Suitability is demonstrated for the solution of the ascent guidance problem by a classical finite difference approach, which can be interpreted as a special form of collocation, but is conceptually simpler and easier to implement.

3) The capability and feasibility of onboard closed-loop ascent guidance are illustrated by a series of carefully designed tests using the data of an RLV.

II. Endoatmospheric Ascent Guidance Problem Formulation

A. Ascent Dynamics

Throughout this paper, the notation is such that a boldface symbol is a vector, and the lightface form of the same symbol indicates the magnitude of this vector. The equations of motion of the RLV in an inertial coordinate system can be expressed as

$$\dot{\boldsymbol{r}} = \boldsymbol{V} \tag{1}$$

$$\dot{\boldsymbol{V}} = \boldsymbol{g}(\boldsymbol{r}) + \boldsymbol{A}/m(t) + T\mathbf{1}_b/m(t) + \boldsymbol{N}/m(t) \tag{2}$$

$$\dot{m} = -(\eta T_{\text{vac}}/g_0 I_{\text{sp}}) \tag{3}$$

where $\boldsymbol{r}$ and $\boldsymbol{V} \in R^3$ are the position and inertial velocity vectors, $\boldsymbol{g}$ is the gravitational acceleration, T_{vac} is the full vacuum thrust magnitude, $\eta > 0$ is the engine throttle, and T is the current thrust magnitude including effects of throttle modulation and thrust loss due to backpressure. In this formulation, the total engine thrust is

Received 3 June 2002; presented as Paper 2002-4558 at the AIAA Guidance, Navigation, and Control Conference, Monterey, CA, 5–8 August 2002; revision received 23 August 2002; accepted for publication 27 August 2002.

*Associate Professor, Department of Aerospace Engineering and Engineering Mechanics; plu@iastate.edu. Associate Fellow AIAA.

†Graduate Research Assistant, Department of Aerospace Engineering and Engineering Mechanics.

‡Graduate Research Assistant, Department of Aerospace Engineering and Engineering Mechanics. Student Member AIAA.

assumed to be aligned with the body longitudinal axis and is not gimbaled independently. The vectors $\boldsymbol{A}$ and $\boldsymbol{N}$ are the aerodynamic forces in the body longitudinal and normal direction, respectively; $\mathbf{1}_b$ is the unit vector defining the RLV body longitudinal axis; and $m(t)$ is the mass of the RLV at the current time t. The specific impulse of the engine is I_{sp} and g_0 represents the gravitational acceleration magnitude on the surface of the Earth. For better numerical conditioning, the following nondimensionalization is used:

1) The distances are normalized by R_0, the radius of the Earth at the equator.

2) Time is normalized by $\sqrt{(R_0/g_0)}$.

3) The velocities are normalized by $\sqrt{(R_0 g_0)}$, the circular velocity around the Earth at R_0.

The gravity is modeled by the Newtonian central gravity field. With some abuse of notation, we use the same names hereafter for the dimensionless variables. The dimensionless equations of motion from Eqs. (1) and (2) are then

$$\boldsymbol{r}' = \boldsymbol{V} \tag{4}$$

$$\boldsymbol{V}' = -(1/r^3)\boldsymbol{r} + \boldsymbol{A} + T\mathbf{1}_b + \boldsymbol{N} \tag{5}$$

where the prime indicates differentiation with respect to the dimensionless time. Now $\boldsymbol{A}$ and $\boldsymbol{N}$ are the aerodynamic accelerations in g_0 in the body longitudinal and normal direction, respectively, and T the magnitude of the thrust acceleration in g_0. The magnitudes of the dimensionless aerodynamic and thrust accelerations are given by

$$A = [R_0\rho_0/2m(t)]\rho(r)V_r^2 S_{\text{ref}} C_A(\text{Mach}, \alpha) \tag{6}$$

$$N = [R_0\rho_0/2m(t)]\rho(r)V_r^2 S_{\text{ref}} C_N(\text{Mach}, \alpha) \tag{7}$$

$$T = \eta[T_{\text{vac}} + \Delta T(r)]/m(t)g_0 \tag{8}$$

where ρ_0 is the atmospheric density at R_0, $\rho(r)$ the atmospheric density at radius r normalized by ρ_0, and V_r is the magnitude of the dimensionless wind-relative velocity

$$\boldsymbol{V}_r = \boldsymbol{V} - \bar{\omega}_E \times \boldsymbol{r} - \boldsymbol{V}_w \tag{9}$$

with $\bar{\omega}_E$ being the Earth angular rotation rate vector normalized by $\sqrt{(g_0/R_0)}$ and $\boldsymbol{V}_w$ the dimensionless wind velocity. The axial and normal aerodynamic coefficients C_A and C_N are functions of angle of attack α and Mach number. They are usually expressed in analytical forms by curve fitting the tabulated data for guidance solutions. The reference area S_{ref} is a constant (and dimensional). The thrust loss inside the atmosphere $\Delta T \leq 0$ is a function of altitude through the dependence of ΔT on ambient pressure. Note that the mass flow rate will be reduced by the same percentage as the thrust when the thrust is throttled down.

Our definitions of vehicle body axis frame $x_b y_b z_b$ follow the zero-sideslip formulation.[7,8] In this formulation, the RLV symmetric plane is assumed to be always the plane formed by the body-axis $\mathbf{1}_b$ and the relative velocity vector $\boldsymbol{V}_r$. Thus, the sideslip angle remains zero. Note that such a body-frame orientation necessitates a roll angle about the longitudinal axis $\mathbf{1}_b$ to null the sideslip in the presence of crosswinds. Physically, this is the so-called fly into the wind maneuver. Thus, the unit vector of the body x axis is the same as $\mathbf{1}_b$; the unit vector of the body y axis is

$$\mathbf{1}_y = \frac{\mathbf{1}_{V_r} \times \mathbf{1}_b}{\left\| \mathbf{1}_{V_r} \times \mathbf{1}_b \right\|} \tag{10}$$

where $\mathbf{1}_{V_r} = \boldsymbol{V}_r / V_r$ is the unit vector in the direction of $\boldsymbol{V}_r$. The unit vector of the body z axis completes the right-hand system $\mathbf{1}_z = \mathbf{1}_b \times \mathbf{1}_y$. Denote the body-normal unit vector by $\mathbf{1}_n = -\mathbf{1}_z$. Then

$$\boldsymbol{A} = -A\mathbf{1}_b, \qquad \boldsymbol{N} = N\mathbf{1}_n \tag{11}$$

Note that the angle of attack is

$$\cos\alpha = \mathbf{1}_b^T \mathbf{1}_{V_r} \qquad \text{or} \qquad |\sin\alpha| = \left\| \mathbf{1}_{V_r} \times \mathbf{1}_b \right\| \tag{12}$$

The following expression for $\mathbf{1}_y$ is preferred to Eq. (10):

$$\mathbf{1}_y = \mathbf{1}_{V_r} \times \mathbf{1}_b \Big/ \sin\alpha \tag{13}$$

This is because this definition is valid for both $\alpha > 0$ and $\alpha < 0$ without causing a 180-deg instantaneous change of direction in $\mathbf{1}_y$ when $\mathbf{1}_{V_r}$ and $\mathbf{1}_b$ cross over each other (α changes sign).

The current conditions $\boldsymbol{r}_0$ and $\boldsymbol{V}_0$ are assumed to be known. The ascent guidance problem is to find the desired body-axis orientation $\mathbf{1}_b(t)$ at each instant that determines the thrust direction and aerodynamic forces during the atmospheric portion of the ascent. The engine throttle η will also need to be determined for enforcing some of the path constraints, as will be discussed later. The final conditions will be the engine-cutoff conditions that ensure insertion into the required orbit. These orbital insertion conditions can in general be written as k, $0 < k \leq 6$, algebraic end conditions

$$\Psi[\boldsymbol{r}(t_f), \boldsymbol{V}(t_f)] = 0, \qquad \Psi \in R^k \tag{14}$$

Specifics on Eq. (14) will also be discussed later. In addition, there will be path constraints on the trajectory for safety and vehicle integrity. The three most common path constraints in ascent guidance will be considered: the product of dynamic pressure and α, axial thrust acceleration, and dynamic pressure,

$$|q\alpha| \leq Q_\alpha \tag{15}$$

$$T \leq T_{\max} \tag{16}$$

$$q \leq q_{\max} \tag{17}$$

where $q = \rho V_r^2/2$. The constants Q_α, $T_{\max}$, and $q_{\max}$ are the respective limits for each of the corresponding constraints. Another common path constraint on $|q\beta|$, the product of dynamic pressure and sideslip angle, is not included here because, in our zero-sideslip formulation, this quantity usually already has small magnitude. The constraints on normal acceleration $|N|$ and angle of attack α, if necessary, can be handled in the same way as constraint (15) is handled. Thus, they will not be discussed separately. Collectively, the given path constraints may be written in a compact form

$$\boldsymbol{S}(\boldsymbol{r}, \boldsymbol{V}, \mathbf{1}_b, t) \leq 0 \tag{18}$$

B. Optimal Control Problem

The mathematical tool used to find the optimal ascent guidance commands is the optimal control theory. In this setting, a performance index is defined. The minimization of this performance index is usually tied in one way or another to the minimization of propellant usage. Denote the performance index by

$$J = \phi(\boldsymbol{r}_f, \boldsymbol{V}_f, t_f) \tag{19}$$

where t_f is the engine cutoff time and $\boldsymbol{r}_f$ and $\boldsymbol{V}_f$ are the position and inertial velocity of the RLV at t_f. The functional form of J is best selected to be most convenient for a particular formulation of the optimal ascent problem. Typical choices are $J = t_f$, for the minimum-time problem, or $J = 1/r_f - V_f^2$, for the maximum-energy problem with a fixed t_f.

Recently, it has been shown that in general no singular optimal thrust programs exist in atmospheric ascent.[9] Therefore, the optimal throttle is bang–bang type. In this paper, the engine throttle η is treated as a given (possibly time-varying) input. Thus, the variation of the mass $m(t)$ is considered a prescribed function of time, not a state. To focus on the presentation of the essentials of the approach, the path constraints (15–17) will be added later. With these assumptions and noting the constraint $\mathbf{1}_b^T \mathbf{1}_b = 1$, we define the Hamiltonian,

$$H = \boldsymbol{p}_r^T \boldsymbol{V} + \boldsymbol{p}_V^T\left[-(1/r^3)\boldsymbol{r} + (T - A)\mathbf{1}_b + N\mathbf{1}_n\right] + \mu\left(\mathbf{1}_b^T \mathbf{1}_b - 1\right) \tag{20}$$

where μ is a scalar multiplier and $\boldsymbol{p}_r$ and $\boldsymbol{p}_V \in R^3$ are the so-called costate vectors. Let an asterisk signify the optimal values of the relevant variables. The standard necessary conditions for the optimal solution are[10]

$$\boldsymbol{p}_r' = -\frac{\partial H}{\partial \boldsymbol{r}} \tag{21}$$

$$\boldsymbol{p}_V' = -\frac{\partial H}{\partial \boldsymbol{V}} \tag{22}$$

$$H(\boldsymbol{p}_r, \boldsymbol{p}_V, \boldsymbol{r}^*, \boldsymbol{V}^*, \mathbf{1}_b^*, t) = \max_{\mathbf{1}_b} H(\boldsymbol{p}_r, \boldsymbol{p}_V, \boldsymbol{r}^*, \boldsymbol{V}^*, \mathbf{1}_b, t) \quad (23)$$

The derivation of the expressions of the costate equations (21) and (22) is quite involved. The detailed equations are provided in Appendix A. The optimal solution must also satisfy the terminal constraints (14) and the following transversality conditions:

$$\boldsymbol{p}_r(t_f) = -\frac{\partial \phi(\boldsymbol{r}_f, \boldsymbol{V}_f, t_f)}{\partial \boldsymbol{r}_f} + \left(\frac{\partial \Psi}{\partial \boldsymbol{r}_f}\right)^T \boldsymbol{\nu} \quad (24)$$

$$\boldsymbol{p}_V(t_f) = -\frac{\partial \phi(\boldsymbol{r}_f, \boldsymbol{V}_f, t_f)}{\partial \boldsymbol{V}_f} + \left(\frac{\partial \Psi}{\partial \boldsymbol{V}_f}\right)^T \boldsymbol{\nu} \quad (25)$$

$$H(\boldsymbol{p}_r, \boldsymbol{p}_V, \boldsymbol{r}^*, \boldsymbol{V}^*, \mathbf{1}_b^*, t)\big|_{t_f} = \frac{\partial \phi}{\partial t_f} \quad (26)$$

where $\boldsymbol{\nu} \in R^k$ is a constant multiplier vector. The last condition (26) is for the cases where the final time t_f is not specified. The first two conditions (24) and (25) can be combined to eliminate the unknown vector $\boldsymbol{\nu}$ and yield $6-k$ independent conditions involving only final costate $\boldsymbol{p}_f = (\boldsymbol{p}_{r_f}^T \boldsymbol{p}_{V_f}^T)^T$ and final state $\boldsymbol{x}_f = (\boldsymbol{r}_f^T \boldsymbol{V}_f^T)^T$. The general approach will be first finding the $6-k$ linear independent solutions of the homogeneous system

$$\left(\frac{\partial \Psi}{\partial \boldsymbol{x}_f}\right)\boldsymbol{\xi} = 0$$

Let $\boldsymbol{\xi}_i(\boldsymbol{x}_f) \in R^6$, $i = 1, \ldots, 6-k$ be such solutions. Note that $\boldsymbol{\xi}_i$ are functions of $\boldsymbol{x}_f$. Transversality conditions (24) and (25) are then equivalent to

$$\left(\boldsymbol{p}_f + \frac{\partial \phi}{\partial \boldsymbol{x}_f}\right)^T \boldsymbol{\xi}_i \triangleq \Gamma_i(\boldsymbol{p}_f, \boldsymbol{x}_f) = 0, \qquad i = 1, \ldots, 6-k \quad (27)$$

The k terminal constraints (14) plus the preceding $6-k$ conditions constitute the six terminal conditions for the optimal control problem. For a given problem, the conditions in Eq. (27) can oftentimes be obtained more conveniently by using the terminal constraints (14) and taking dot products of Eqs. (24) and (25) with appropriate vectors related to the final state $\boldsymbol{x}_f$. Examples will be given later.

The optimality condition (23) necessitates

$$\frac{\partial H}{\partial \mathbf{1}_b} = 0 \quad (28)$$

Define $s = \|\mathbf{1}_b \times \mathbf{1}_{V_r}\|$. The expansion of the optimality condition $\partial H/\partial \mathbf{1}_b = 0$ requires, among other relationships, the following:

$$\frac{\partial \alpha}{\mathbf{1}_b} = \frac{\cos\alpha}{\sin\alpha}\mathbf{1}_b - \frac{1}{\sin\alpha}\mathbf{1}_{V_r} \quad (29)$$

$$\frac{\partial \mathbf{1}_n}{\partial \mathbf{1}_b} = \frac{1}{s}\left\{\left(\mathbf{1}_{V_r}^T \mathbf{1}_b\right) I_3 + \mathbf{1}_b \mathbf{1}_{V_r}^T + \frac{1}{s^2}\left[\left(\mathbf{1}_{V_r}^T \mathbf{1}_b\right)\mathbf{1}_b - \mathbf{1}_{V_r}\right]\left[\left(\mathbf{1}_{V_r}^T \mathbf{1}_b\right)\mathbf{1}_{V_r} - \mathbf{1}_b\right]^T\right\} \quad (30)$$

where I_3 is a 3×3 identity matrix. Let

$$a = p_V\left[\left(\mathbf{1}_{V_r}^T \mathbf{1}_b\right)\left(\mathbf{1}_{p_V}^T \mathbf{1}_b\right) - \left(\mathbf{1}_{V_r}^T \mathbf{1}_{p_V}\right)\right]/s$$

$$b = -p_V A_\alpha + a N_\alpha$$

where $\mathbf{1}_{p_V} = \boldsymbol{p}_V / p_V$, $A_\alpha = \partial A/\partial\alpha$, and $N_\alpha = \partial N/\partial\alpha$. When $\partial H/\partial \mathbf{1}_b$ is evaluated, keep in mind that both A and N are functions of α, therefore, functions of $\mathbf{1}_b$. Carrying out the differentiations and collecting terms, we eventually have

$$\mathbf{1}_b^* = [1/(2\mu + b/\tan\alpha - aN/s^2)]\left\{-\left[T - A + N\left(\mathbf{1}_{V_r}^T \mathbf{1}_b^*\right)/s\right]\boldsymbol{p}_V + \left[b/\sin\alpha - N p_V\left(\mathbf{1}_{p_V}^T \mathbf{1}_b^*\right)/s - aN\left(\mathbf{1}_{V_r}^T \mathbf{1}_b^*\right)/s^2\right]\mathbf{1}_{V_r}\right\}$$

$$\triangleq c_1(\boldsymbol{x}, \boldsymbol{p}, \mathbf{1}_b^*)\boldsymbol{p}_V + c_2(\boldsymbol{x}, \boldsymbol{p}, \mathbf{1}_b^*)\boldsymbol{V}_r \quad (31)$$

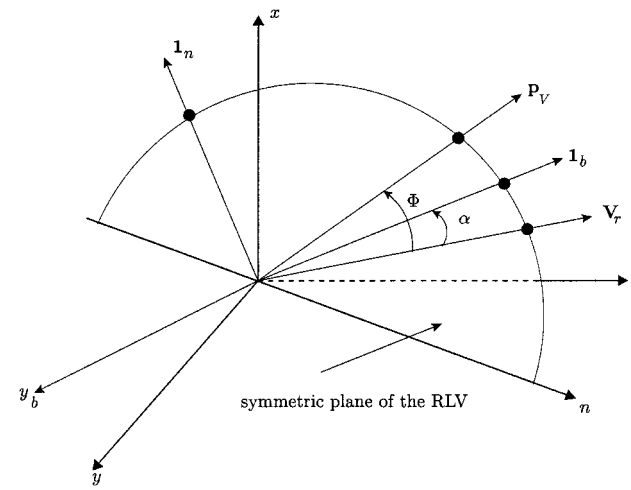

Fig. 1 Body axes, primer vector $\boldsymbol{p}_V$, and $\boldsymbol{V}_r$ in launch plumbline system *xyz*.

where c_1 and c_2 are scalar functions of the state, costate, and $\mathbf{1}_b^*$. Hence, we conclude that the optimal body axis lies in the plane formed by the primer vector $\boldsymbol{p}_V$ and relative velocity vector $\boldsymbol{V}_r$. A similar conclusion is reached by using a geometric approach in an earlier work by Vinh,[11] where the thrust direction and aerodynamic force vector are assumed to be independent controls.

Condition (31) suggests that the search for the optimal body-axis orientation can be reduced to a one-dimensional search in the plane of $\boldsymbol{p}_V$ and $\boldsymbol{V}_r$ (Ref. 7). Let Φ be the angle between the vectors $\boldsymbol{p}_V$ and $\boldsymbol{V}_r$. At each instant with given state and costate, Φ is known. Denote by $\mathbf{1}_{p_r}$ and $\mathbf{1}_{p_V}$ the unit vectors in the directions of $\boldsymbol{p}_r$ and $\boldsymbol{p}_V$, respectively. Then it is straightforward to see from Fig. 1 that $\mathbf{1}_b^T \mathbf{1}_{p_V} = \cos(\Phi - \alpha)$ and $\mathbf{1}_n^T \mathbf{1}_{p_V} = \sin(\Phi - \alpha)$. When these two relationships are used in the expression of H in Eq. (20), it is clear that maximizing H with respect to $\mathbf{1}_b$ is equivalent to $\partial H/\partial\alpha = 0$, which in turn results in[7]

$$\tan(\Phi - \alpha)(T - A + N_\alpha) - (A_\alpha + N) = 0 \quad (32)$$

Because A, N, A_α, and N_α are generally functions of α, Eq. (32) needs to be solved numerically for α. An interesting observation on the role of the primer vector $\boldsymbol{p}_V$ in atmospheric ascent can be made based on Eq. (32), when $A_\alpha = N_\alpha = 0$. Assume the following:

1) The thrust vector remains in the symmetrical plane of the launch vehicle, but otherwise the thrust can be aligned with the longitudinal body axis, canted, or independently gimbaled.

2) The normal and axial aerodynamic forces are independent of α.

3) The sideslip angle is zero.

Then the optimal total force of the thrust and aerodynamic forces is in the direction of the primer vector $\boldsymbol{p}_V$.

The cases for canted or gimbaled thrust vector are not included in the preceding formulation. However, the same conclusion can be reached in a similar fashion as in preceding analysis. This result is an extension of the well-known primer vector theory on optimal rocket flight in vacuum.[12]

Once α is found from Eq. (32), c_1 and c_2 in condition (31) can be solved in terms of α and Φ by taking the dot product of condition (31) with $\mathbf{1}_{p_V}$ and with $\mathbf{1}_{V_r}$:

$$\mathbf{1}_b^* = \left(\frac{\sin\alpha}{\sin\Phi}\right)\mathbf{1}_{p_V} + \left[\frac{\cos\alpha - \cos\Phi\cos(\Phi - \alpha)}{\sin^2\Phi}\right]\mathbf{1}_{V_r} \quad (33)$$

Note that as the atmospheric density decreases (approaching vacuum flight) the aerodynamic terms diminish and that $\alpha \to \Phi$ from Eq. (32). The optimal body axis in Eq. (33) and, therefore, the optimal thrust vector become aligned with the primer vector $\boldsymbol{p}_V$ (Ref. 12).

The case where coast arcs are allowed within a stage or between stages will not be discussed in this paper. The reader is referred to Gath and Calise[7] for a method to determine the optimal lengths of coast arcs.

C. Determination of Correct Signs of Φ and α

We stress that for onboard guidance applications, particularly for nonaxisymmetric RLV, it is imperative that great care be taken to determine the sign of Φ in Eqs. (32) and (33). Not only the value and sign of α depend on the sign of Φ, but more important is the physical implication of the sign change of Φ. The optimal body y axis, shown as y_b in Fig. 1, can also be defined by

$$\mathbf{1}_y = \frac{\mathbf{1}_{V_r} \times \mathbf{1}_{p_V}}{\sin\Phi} \tag{34}$$

When the vector $\mathbf{1}_{V_r} \times \mathbf{1}_{p_V}$ changes direction by 180 deg, the sign of Φ should change accordingly. If the correct sign of Φ is not identified when it changes, the body axis y_b will have an instantaneous change of direction by 180 deg. Consequently, the guidance system would command an unnecessary 180-deg roll of the vehicle. For axisymmetric launch vehicles, this may not be a concern because of the symmetry of aerodynamic properties with respect to the angle of attack regardless of its sign. Steps can be taken to recognize and ignore such large roll commands without affecting the guidance precision. However, for a nonaxisymmetric RLV, the determination of the correct sign of Φ, hence, the sign of α, is a very important issue.

Considerable effort has been spent on finding a reliable way to determine the sign of Φ. This task turns out to be not nearly as straightforwardas it might appear. Because the RLV is going through three-dimensionalflight with pitch, yaw, and roll motion, there is no clear-cut reference that can be used to determine if $\boldsymbol{p}_V$ is "above" $\boldsymbol{V}_r(\Phi > 0)$ or "below" $\boldsymbol{V}_r(\Phi < 0)$. What is described next is the simplest among several techniques we have developed and yet appears to be the one that works the best so far.

Define the inertial launch plumbline coordinate system xyz as follows: The origin is at the center of the Earth; the x axis is parallel to the gravity vector at the launch site at the time of launch, positive up; and the z axis is parallel to and positive in the same direction as the chosen Earth-fixed launch azimuth. The launch azimuth angle A_z is calculated by

$$\sin A_z = \cos i / \cos\phi_c \tag{35}$$

where i is the target orbit inclination and ϕ_c the geocentric latitude of the launch site. Then the y axis of the launch plumbline system completes a right-hand system. Figure 1 shows the vectors $\boldsymbol{p}_V$, $\boldsymbol{V}_r$, $\mathbf{1}_b$, and $\mathbf{1}_n$ in the launch plumbline coordinate system. The RLV symmetric plane is the same plane formed by $\boldsymbol{p}_V$ and $\boldsymbol{V}_r$ as shown in the preceding section. Define the vector

$$\boldsymbol{n} = \boldsymbol{i} \times (\boldsymbol{V}_r \times \boldsymbol{p}_V) \tag{36}$$

where $\boldsymbol{i}$ is the unit vector in the plumbline x-axis direction. The vector $\boldsymbol{n}$, also shown in Fig. 1, is analogous to the nodal vector of an orbit in the xyz system, if the RLV symmetric plane is regarded as the orbital plane. Assume that the RLV body axes coincide with the launch plumbline axes at launch, as is usually the case. The magnitude of Φ is found from $\cos\Phi = \mathbf{1}_{p_V}^T \mathbf{1}_{V_r}$. In the standard yaw–pitch–roll rotation sequence of the Euler angles, when the yaw and roll angle are within and not near to ± 90 deg, the sign of Φ is determined by

$$\begin{aligned} &\Phi > 0, \quad \text{if} \quad n_z > 0 \\ &\Phi < 0, \quad \text{if} \quad n_z < 0 \end{aligned} \tag{37}$$

where n_z is the z component of the vector $\boldsymbol{n}$ in the launch plumbline system. This criterion is again analogous to the one used to determine whether an orbit is posigrade or retrograde.

D. Adding the Path Constraints

When inequality constraints (18) are added to the problem, the costate equations (21) and (22) will have additional terms related to the constraints, and the condition to determine the controls will change. We discuss each of the three constraints in Eqs. (15–17) as follows.

Constraint $S_1 = q\alpha - Q_\alpha \le 0$

Without loss of generality, the absolute sign in Eq. (15) is removed for simplicity of discussion. This constraint is a zeroth-order constraint in that the control $\mathbf{1}_b$ appears explicitly (through α) in the constraint itself. In such a case, the costate equations take the form

$$\boldsymbol{p}' = -\frac{\partial H}{\partial \boldsymbol{x}} - \lambda_{q\alpha}\frac{\partial S_1}{\partial \boldsymbol{x}} \tag{38}$$

where the multiplier $\lambda_{q\alpha} = 0$ and the problem is the same as in the preceding sections when $S_1 < 0$. When $S_1 = 0$, then $\lambda_{q\alpha}$ satisfies the condition

$$\frac{\partial H}{\partial \mathbf{1}_b} + \lambda_{q\alpha}\frac{\partial S_1}{\partial \mathbf{1}_b} = 0 \tag{39}$$

Using condition (29), we have

$$\frac{\partial S_1}{\partial \mathbf{1}_b} = q\frac{\partial \alpha}{\mathbf{1}_b} = q\left(\frac{\cos\alpha}{\sin\alpha}\mathbf{1}_b - \frac{1}{\sin\alpha}\mathbf{1}_{V_r}\right) \tag{40}$$

Using condition (39) and the result in Eq. (40) and following the steps similar to those in Sec. II.B, we can show that the optimal body axis $\mathbf{1}_b^*$ is again in the plane of $\boldsymbol{p}_V$ and $\boldsymbol{V}_r$ as before. Note that this conclusion cannot be assumed to be always true for any inequality constraints and must be verified for each case. In this case, condition (39) is equivalent to

$$\frac{\partial H}{\partial \alpha} + \lambda_{q\alpha}\frac{\partial S_1}{\partial \alpha} = 0 \tag{41}$$

or

$$\lambda_{q\alpha} = -\frac{\partial H/\partial\alpha}{q} \tag{42}$$

where

$$\frac{\partial H}{\partial \alpha} = p_V[(T - A + N_\alpha)\sin(\Phi - \alpha) - (A_\alpha + N)\cos(\Phi - \alpha)] \tag{43}$$

In a finite interval where $S_1 = 0$, the angle of attack is directly obtained from

$$\alpha = Q_\alpha / q \tag{44}$$

The direction of the body axis $\mathbf{1}_b$ is determined by α in the plane of $\boldsymbol{p}_V$ and $\boldsymbol{V}_r$ by Eq. (33) as before. The multiplier $\lambda_{q\alpha}$ is calculated from Eq. (42) and used in the costate equation (38) to propagate the costate.

Constraint $S_2 = T - T_{\max} \le 0$

Typically for the usual values of $T_{\max}$, this constraint only becomes active after the trajectory is outside the dense atmosphere where $\Delta T(r) \approx 0$. Thus, whether or not and when this constraint will become active are uniquely determined by the prescribed engine throttle, not influenced by any of the trajectory state variables and the control $\mathbf{1}_b$. In this sense it is not a state or control constraint. (Recall that the engine throttle is regarded as a prescribed input.) Once $S_2 = 0$, the engine throttle will be adjusted according to

$$\eta = T_{\max} m(t) g_0 / T_{\text{vac}} \tag{45}$$

to keep $S_2 = 0$. The costate equations are the same as in Eqs. (21) and (22). and the optimal body axis $\mathbf{1}_b^*$ is found in Eq. (33) as before.

Constraint $S_3 = q - q_{\max} \le 0$

This is a first-order constraint because the control $\mathbf{1}_b$ appears in the first-order derivative of S_3. Recall that $q = \rho(r)V_r^2/2$, and $\boldsymbol{V}_r$ is defined in Eq. (9). Thus,

$$S_3' = (1/2r)\rho_r V_r^2 \boldsymbol{r}^T \boldsymbol{V} + \rho \boldsymbol{V}_r^T \boldsymbol{V}_r' \tag{46}$$

where $\rho_r = \partial\rho/\partial r$ and

$$\boldsymbol{V}_r' = \boldsymbol{V}' - \bar{\omega}_E \times \boldsymbol{V} - \boldsymbol{V}_w' \tag{47}$$

Suppose that $S_3 = 0$ in a finite interval $[t_1, t_2]$. The costate equations in this interval are

$$\boldsymbol{p}' = -\frac{\partial H}{\partial \boldsymbol{x}} - \lambda_q \frac{\partial S_3'}{\partial \boldsymbol{x}} \tag{48}$$

The corresponding optimality condition is

$$\frac{\partial H}{\partial \mathbf{1}_b} + \lambda_q \frac{\partial S_3'}{\partial \mathbf{1}_b} = 0 \tag{49}$$

The costate will have a jump at t_1,

$$\boldsymbol{p}\left(t_1^+\right) = \boldsymbol{p}\left(t_1^-\right) + \kappa \frac{\partial S_3}{\partial \boldsymbol{x}} \tag{50}$$

where κ is a constant multiplier. It can be shown that

$$\frac{\partial S_3'}{\partial \mathbf{1}_b} = d_1(\boldsymbol{x}, \mathbf{1}_b)\boldsymbol{V}_r + d_2(\boldsymbol{x}, \mathbf{1}_b)\mathbf{1}_b \tag{51}$$

where d_1 and d_2 are two scalar functions. Therefore, using the derivations in Sec. II.B, we can show that condition (49) still results in that the optimal body axis $\mathbf{1}_b^*$ lies in the plane of $\boldsymbol{p}_V$ and $\boldsymbol{V}_r$. Conditions similar to Eqs. (41) and (42) can now be readily derived. And the condition $S_3' = 0$ provides the equation for the solution of α in $[t_1, t_2]$.

With the preceding equations laid out, the optimal solution in principle can be found numerically. However, two implementation issues arise: 1) The engine throttle is known to be more effective in regulating the dynamic pressure by slowing down the increase of the velocity. 2) The jump condition Eq. (50) makes it necessary to estimate accurately the time t_1 in the solution process for it to converge quickly. The second issue may not be a problem if only nominal ascent is considered because a good estimate of t_1 can be obtained offline if the vehicle modeling, day-of-launch wind data, and mission parameters are well known. However, a chief potential advantage of closed-loop ascent guidance is for autonomous abort guidance. In aborts, the conditions are inevitably well off nominal, and any offline estimates of t_1 could be no better than an arbitrary guess.

In an onboard environment, the entire ascent trajectory $\boldsymbol{x}(\cdot)$, along with the control $\mathbf{1}_b(\cdot)$ and throttle $\eta(\cdot)$, is generated from the current condition $\boldsymbol{x}(t)$ to the target condition in each guidance cycle. The current attitude commands and throttle command are from $\mathbf{1}_b(t)$ and $\eta(t)$, the first data point in the guidance solution. To address the issues and keep the algorithms reasonably simple and robust for onboard guidance purposes, we adopt the following approach: The optimal body axis in the guidance solution is still determined as in Secs. II.B and II.C with the prescribed throttle where no constraint on the dynamics pressure is considered. This is justified because we conclude that the optimal body axis remains in the plane of $\boldsymbol{p}_V$ and $\boldsymbol{V}_r$ even when $q \le q_{\max}$ is active. Next, we consider the first-order derivative of q at t that can be obtained by combining Eqs. (5), (8), (46), and (47):

$$q'(t) = a_q(\boldsymbol{x}, \mathbf{1}_b)\eta(t) + b_q(\boldsymbol{x}, \mathbf{1}_b) \tag{52}$$

where η is the current engine throttle and the expressions for a_q and b_q are readily available when one makes the substitutions. Let $\delta > 0$ be a time increment. A first-order approximation of $q(t+\delta)$ is

$$q(t+\delta) \approx q(t) + q'(t)\delta = q(t) + [a_q(\boldsymbol{x}, \mathbf{1}_b)\eta(t) + b_q(\boldsymbol{x}, \mathbf{1}_b)]\delta \tag{53}$$

When deciding the throttle command at the current time t, we require that $q(t+\delta) \le q_{\max}$. Using the preceding approximation for $q(t+\delta)$ in this condition, we have

$$\eta(t) \le \frac{q_{\max} - q(t) - b_q\delta}{a_q\delta} \triangleq \eta_q \tag{54}$$

Most likely η will have a minimum allowable setting $\eta_{\min} > 0$ during engine-on period. Let $\eta_{\text{prb}} > \eta_{\min}$ be the otherwise prescribed throttle setting, for example, $\eta_{\max}$. The current engine throttle command is determined by

$$\eta = \begin{cases} \eta_{\text{prb}}, & \text{if} \quad \eta_q > \eta_{\text{prb}} \\ \eta_q, & \text{if} \quad \eta_{\min} \le \eta_q \le \eta_{\text{prb}} \\ \eta_{\min}, & \text{if} \quad \eta_q < \eta_{\min} \end{cases} \tag{55}$$

The last case in Eq. (55) would be when the dynamic pressure constraint cannot be met by lowering the engine throttle within the allowable range, which is an unlikely event.

Remarks:

1) Because a first-order expansion (53) is used in deriving condition (54), one may be tempted to suspect that the enforcement of $q \le q_{\max}$ by this approach is acceptable only when δ is small. On the contrary, it can be shown that as long as the trajectory begins in the region where $q < q_{\max}$, the throttle given by Eq. (55) guarantees strict satisfaction of the constraint $q \le q_{\max}$ for any $\delta > 0$. It turns out that when $q \to q_{\max}$, $\eta = \eta_q$ drives $q' \to 0$ for any $\delta > 0$, and whenever $q = q_{\max}$, $q' = 0$. A detailed discussion of strict enforcement of first-order constraints by this type of technique can be found in a previous work on constrained nonlinear control systems.[13]

2) Because of the property just described, an appropriate value δ can always be chosen for a given vehicle to eliminate the undesirable jitters in throttle command when the dynamic pressure constraint is active and still accurately enforce the constraint. Typically there is a minimum value of δ above which a δ will render the engine throttle commands from Eq. (55) to be sufficiently smooth.

3) The benefits of this approach for enforcing the constraint $q \le q_{\max}$ are its simplicity, robustness, and effectiveness. There is no need for the estimate of the instant t_1 when the constraint becomes active. It is closed loop in nature, unlike the open-loop "throttle bucket" approach currently in use for the shuttle. Note, however, that the adjustment of the throttle is done outside the solution process of the optimal control problem. Thus, the guidance solution in the period when the dynamic pressure constraint is active is not strictly optimal in theory (although the body axis is still determined according to the necessary conditions of the optimal control problem). However, this period is always relatively short compared to the total burn time. The impact of this short deviation from the theoretical optimal conditions on performance appears to be negligible. We have used two different state-of-the-art trajectory optimization software packages[14,15] to cross check the performance of the trajectories obtained herein. The differences in terms of propellant consumption have been minimal. (See numerical results in Sec. IV.B). A similar observation is also made by Corvin[16] on the effects of adjusting throttle to control dynamic pressure.

III. Numerical Method

A. Finite Difference Approach

Finite difference is one of several classical techniques used for two-point boundary-value problems (TPBVPs).[17] For ascent guidance applications, we have found that this classical approach is well suited for the problem.

After the substitution of the control $\mathbf{1}_b$ of Eq. (33) in the equations of motion and costate equations, and the denotation of $\boldsymbol{y} = (\boldsymbol{x}^T \boldsymbol{p}^T)^T \in R^{2n}$ with $n = 6$, the complete TPBVP is now

$$\boldsymbol{y}' = \boldsymbol{f}(t, \boldsymbol{y}) \tag{56}$$

$$\boldsymbol{B}_0(\boldsymbol{y}_0) = 0 \tag{57}$$

$$\boldsymbol{B}_f(\boldsymbol{y}_f) = 0 \tag{58}$$

where $\boldsymbol{B}_0 = \boldsymbol{x}(t_0) - \boldsymbol{x}_0$ in our problem represents the given initial condition and $\boldsymbol{B}_f(\boldsymbol{y}_f) = 0$ are the six final conditions from combining the orbital insertion conditions (14) and the transversality conditions (27). Let t_f be the specified final time. The TPBVP is to find a solution $\boldsymbol{y}(t)$ that satisfies the differential equations (56) and boundary conditions (57) and (58). To find the solution, divide the time interval $t_f - t_0$ into M subintervals of the same length $h = (t_f - t_0)/M$. Let $\boldsymbol{y}_k = \boldsymbol{y}(t_0 + kh)$ be the value of the solution at the node $t_k = t_0 + kh$, $k = 0, \ldots, M$. At the middle point between

t_{k-1} and t_k, denoted by $t_{k-1/2} = t_k - h/2$, the differential equations (56) are approximated by central finite difference:

$$1/h(\boldsymbol{y}_k - \boldsymbol{y}_{k-1}) = \boldsymbol{f}[t_{k-1/2}, (\boldsymbol{y}_k - \boldsymbol{y}_{k-1})/2] \tag{59}$$

or, equivalently, $\boldsymbol{y}_k$ and $\boldsymbol{y}_{k-1}$ are constrained by the equation

$$\boldsymbol{E}_k(\boldsymbol{y}_k, \boldsymbol{y}_{k-1}) \triangleq \boldsymbol{y}_k - \boldsymbol{y}_{k-1} - h\boldsymbol{f}[(t_{k-1/2}, (\boldsymbol{y}_k - \boldsymbol{y}_{k-1})/2] = 0$$
$$k = 1, \ldots, M-1 \tag{60}$$

In addition, the boundary conditions are denoted by

$$\boldsymbol{E}_0(\boldsymbol{y}_0) = \boldsymbol{B}_0(\boldsymbol{y}_0) = 0 \tag{61}$$

$$\boldsymbol{E}_M(\boldsymbol{y}_M) = \boldsymbol{B}_f(\boldsymbol{y}_f) = 0 \tag{62}$$

Treat $\boldsymbol{Y} = (\boldsymbol{y}_0^T \boldsymbol{y}_1^T, \ldots, \boldsymbol{y}_M^T)^T \in R^{2n(M+1)}$ as the unknowns. The same number of equations are

$$\boldsymbol{E}(\boldsymbol{Y}) = 0 \tag{63}$$

where $\boldsymbol{E} = (\boldsymbol{E}_0^T \boldsymbol{E}_1^T, \ldots, \boldsymbol{E}_M^T)^T$. Now the problem becomes a root-finding problem for a system of $2n(M+1)$ nonlinear algebraic equations (63). It has been rigorously established that under certain conditions on smoothness and the boundary conditions, the following holds true[17]:

1) Both the original TPBVP and the finite difference problem have a unique solution.

2) The solution of the preceding finite difference problem $\boldsymbol{y}_k$ is a second-order approximation to the solution of the TPBVP $\boldsymbol{y}^*(t)$ at each t_k, that is,

$$\left\| \boldsymbol{y}^*(t_k) - \boldsymbol{y}_k \right\| = \mathcal{O}(h^2), \qquad k = 1, 2, \ldots, M \tag{64}$$

where $\lim_{h \to 0} \mathcal{O}(h^2)/h^2 < \infty$.

For ascent guidance applications, because the time to go $t_f - t_0$ is decreasing, the accuracy of the finite difference solution will be higher and higher as h becomes smaller even for a small-to-moderate number of nodes.

B. Update Algorithm

The modified Newton method (see Ref. 18) is probably the most suitable algorithm for solving problem (63). When started from an initial guess $\boldsymbol{Y}_0$, the search direction $\boldsymbol{d}_j$ in the jth iteration is determined by solving the linear algebraic equations

$$\left[\frac{\partial \boldsymbol{E}(\boldsymbol{Y}_{j-1})}{\partial \boldsymbol{Y}}\right] \boldsymbol{d}_j = -\boldsymbol{E}(\boldsymbol{Y}_{j-1}), \qquad j = 1, 2, \ldots \tag{65}$$

Next, determine the step size parameter σ_j by the following criterion:

$$\sigma_j = \max_{0 \le i} \left\{ 1/2^i \,\middle|\, \boldsymbol{E}^T\left[\boldsymbol{Y}_{j-1} + (1/2^i)\boldsymbol{d}_j\right]\boldsymbol{E}\left[\boldsymbol{Y}_{j-1} + (1/2^i)\boldsymbol{d}_j\right] < \boldsymbol{E}^T(\boldsymbol{Y}_{j-1})\boldsymbol{E}(\boldsymbol{Y}_{j-1}) \right\} \tag{66}$$

In other words, starting from $\sigma_j = 1$, parameter σ_j is halved repeatedly if necessary until the preceding condition is satisfied. Then the update is given by

$$\boldsymbol{Y}_j = \boldsymbol{Y}_{j-1} + \sigma_j \boldsymbol{d}_j, \qquad 0 < \sigma_j \le 1 \tag{67}$$

This choice of the search step size ensures that the sequence $\{\|\boldsymbol{E}(\boldsymbol{Y}_j)\|\}$ is monotonically decreasing. Convergence is achieved when $\|\boldsymbol{E}(\boldsymbol{Y}_j)\|$ is no greater than a preselected tolerance. The possible additional function evaluations required in checking the preceding step-size condition pose a negligible computational burden because function evaluations are not expensive in this setting. The result, on the other hand, is a much more robust algorithm, especially when the initial guesses are not close to the final solution. The step-size selection (66) is a deciding factor for the success of the finite difference approach in solving the optimal ascent problem.

The evaluation of the Jacobian $\partial \boldsymbol{E}/\partial \boldsymbol{Y}$ can certainly be done analytically. However, we believe that simple numerical finite differencing is more advantageous in this case. This is because, again, unlike in the cases where integrations of differential equations are involved for each function evaluation, the function evaluations here are purely algebraic and fast. Using analytical Jacobian offers no clear computational speed benefits. In our comparison study, we have seen that the numerical Jacobians and analytical Jacobians match between the sixth and eighth digit. With the scalings described in Sec. II.A, this type of precision appears to be more than adequate for convergence to occur. On the other hand, analytical Jacobian will make the computer code significantly more complicated because second-order partial derivatives of the right-hand sides of the state equations are needed. Also, when some of the path constraints in Eq. (18) become active, the associated Lagrange multipliers such as the one in Eq. (42) will be functions of state and costate, adding more complexity to the analytical Jacobian. When the RLV design is evolving in the developmental stage, or if the same ascent guidance code is desired to be applied to different RLV configurations, fewer labor-intensive changes to the code would be required with numerical differencing. One exception is for the gradients of the boundary conditions in Eq. (62). The gradients for these conditions are evaluated analytically because they are readily available.

At first glance, solving the linear system (65) may seem to be a formidable task in an onboard environment because of the dimension of the problem $[2n(M+1)]$, which can easily be over 1000. However, a closer inspection reveals that the Jacobian matrix in system (65) has a special sparse pattern due to the dependence of $\boldsymbol{E}_k$ on only $\boldsymbol{y}_k$ and $\boldsymbol{y}_{k-1}$ [see Eq. (60)]. Therefore, a very efficient algorithm, both in speed and storage, based on Gauss eliminations and sequential backsubstitutions can be devised to solve system (65). Details in implementation of such an algorithm are well documented.[19]

C. Continuation on Atmospheric Density and Initial Guesses

It is well known in ascent trajectory optimization that the strong coupling of the aerodynamic forces with the orientation of the body axis and the inequality path constraints make the convergence of any algorithm from a completely "cold" initial start difficult to achieve. That is the reason why homotopic filters,[6] sometimes up to four levels of homotopy,[7] are used to distort the solution gradually from the vacuum solution to the final solution. A similar approach is taken here, except that the finite difference approach described in the preceding section appears to need only one level of homotopy. The form of the continuation used here is simpler and is only on the atmospheric density:

$$\hat{\rho} = \varepsilon \rho, \qquad 0 \le \varepsilon \le 1 \tag{68}$$

For each fixed ε, the parameter $\hat{\rho}$ is used in place of the atmospheric density everywhere ρ appears in the state equations, costate equations, and path constraints. Note that all of the development in this paper, although aimed at atmospheric ascent, applies without any modifications to vacuum flight when $\rho = 0$. The homotopic parameter ε is initiated at 0 for the vacuum solution and gradually increased to unity for full atmospheric solution. The converged solution for the current value of ε serves as the starting point for the solution with the next value of ε till $\varepsilon = 1$. We have been using an increment of 0.1 on ε, which seems to be adequate in our cases. In onboard applications, no continuation is usually necessary when the solution in the previous guidance cycle is used as the starting point.

For the vacuum solution, in many cases linear interpolations between the initial values and targeted final values of the state and constant guesses for $\boldsymbol{p}(t)$ suffice to be an initial guess $\boldsymbol{Y}_0$ that will lead to convergence. An analytical method to generate optimal vacuum ascent trajectories is presented in Appendix B. This method is based on a succinct summary by Calise et al.[6] of a number of elegant results on optimal vacuum guidance developed in the last three decades. Some modifications have been made in the version in Appendix B. With this type of algorithm, an optimal vacuum solution can be found very quickly and reliably. This solution can in turn be used as the starting solution in the preceding continuation process.

IV. Testing

A number of test cases are presented in this section. All of the cases use the vehicle parameters, mass property, propulsion system

modeling, and aerodynamic modeling of the X-33 vehicle. The X-33 is a single-stage suborbital vehicle with a lifting-body configuration. To make the vehicle orbit capable, the specific impulse is doubled in the simulations. All of the missions are launched from the NASA Kennedy Space Center (KSC). The open-loop solutions obtained using the method described in this paper are compared with those obtained by using other methods. Closed-loop simulations are performed to assess the feasibility of onboard ascent guidance with the approach proposed in this paper. Unless stated otherwise, all of the vectors are expressed in the inertial launch plumbline coordinate system introduced in Sec. II.C.

A. Terminal Conditions and Final Time Adjustment

The finite difference approach described in Sec. III.A is most convenient for fixed final-time problems. Similar to what is done in Ref. 7, the optimal ascent problem is solved as a series of fixed final-time problems to maximize the orbital energy, that is, to minimize

$$J = 1/r_f - V_f^2/2 \tag{69}$$

The final time is adjusted sequentially until the optimal value of J is equal to the specified (negative) orbital energy.

Four-Constraint Problem

The orbital insertion conditions are given by the radius r_f^*, velocity V_f^*, orbital inclination i^*, and flight-path angle γ_f^*. Note that γ_f^* need not necessarily be zero in this case. These conditions are equivalent to specified semimajor axis, eccentricity, orbital inclination, and true anomaly at insertion point. Let $\mathbf{1}_N$ be a unit vector parallel to the polar axis of the Earth and pointing to the north. For each fixed t_f the three orbital insertion conditions (14) in this case can be expressed as

$$\tfrac{1}{2}\boldsymbol{r}_f^T\boldsymbol{r}_f - \tfrac{1}{2}r_f^{*2} = 0 \tag{70}$$

$$\mathbf{1}_N^T(\boldsymbol{r}_f \times \boldsymbol{V}_f) - \|\boldsymbol{r}_f \times \boldsymbol{V}_f\| \cos i^* = 0 \tag{71}$$

$$\boldsymbol{r}_f^T\boldsymbol{V}_f - r_f V_f \sin\gamma_f^* = 0 \tag{72}$$

For simplicity, we replace r_f in Eq. (69) with r_f^*. Taking dot products of the transversality condition Eq. (24) with $\boldsymbol{r}_f$, Eq. (25) with $\boldsymbol{r}_f$, Eq. (24) with $\boldsymbol{V}_f$, and both Eqs. (24) and (25) with $\boldsymbol{h}_f = \boldsymbol{r}_f \times \boldsymbol{V}_f$, we can eliminate the multiplier vector $\boldsymbol{\nu}$ and obtain

$$\left(\boldsymbol{V}_f^T\boldsymbol{p}_{r_f}\right)r_f^2 - \left(\boldsymbol{r}_f^T\boldsymbol{p}_{V_f}\right)V_f^2 + \left(\boldsymbol{r}_f^T\boldsymbol{V}_f\right)\left(V_f^2 - \boldsymbol{r}_f^T\boldsymbol{p}_{r_f}\right) = 0 \tag{73}$$

$$\boldsymbol{V}_f^T\boldsymbol{p}_{V_f} - V_f^2 = 0 \tag{74}$$

$$\left(\boldsymbol{h}_f^T\boldsymbol{p}_{r_f}\right)\left[\boldsymbol{h}_f^T(\boldsymbol{r}_f \times \mathbf{1}_N)\right] + \left(\boldsymbol{h}_f^T\boldsymbol{p}_{V_f}\right)\left[\boldsymbol{h}_f^T(\boldsymbol{V}_f \times \mathbf{1}_N)\right] = 0 \tag{75}$$

Equations (70–75) constitute the terminal boundary conditions (58) for this case.

To adjust the final time, we have found that the secant method works very well for this purpose. Let $t_f^{(k-1)}$ and $t_f^{(k)}$ be two consecutive estimates of t_f used to solve the given problem and $J^{(k)}$ and $J^{(k-1)}$ be the values of J when $t_f^{(k)}$ and $t^{(k-1)}$ are used to solve the optimal ascent problem. Then the next choice of t_f is given by

$$t_f^{(k+1)} = t_f^{(k)} - \left(\frac{t_f^{(k)} - t_f^{(k-1)}}{J^{(k)} - J^{(k-1)}}\right)(J^{(k)} - J^*), \qquad k = 1, 2, \ldots \tag{76}$$

The correct t_f is found when $|J^{(k)} - J^*|$ is within a preset tolerance, and the final velocity V_f will be within a small neighborhood of V_f^*. To use the secant search (76), two starting values of $t_f^{(0)}$ and $t_f^{(1)}$ are needed. For onboard guidance, the converged value of t_f in the previous guidance cycle is the logical choice of $t_f^{(0)}$. Because the guidance solution in the current cycle will be very close to the previous one, the variational relationship of $\delta J = H(t_f)\delta t_f$ in optimal control problems can be used to generate $t_f^{(1)}$,

$$t_f^{(1)} = t_f^{(0)} + \left[(J^{(0)} - J^*)/H\left(t_f^{(0)}\right)\right] \tag{77}$$

where $H(t_f^{(0)})$ is the final value of the Hamiltonian. It is possible to simply use Eq. (77) to search for all $t_f^{(k)}$, $k \geq 1$, as in Refs. 7 and 8. However, we have found the secant method (76) to be much more robust and efficient than the exclusive use of Eq. (77).

Five-Constraint Problem

The orbital insertion conditions are given by the radius r_f^*, velocity V_f^*, orbital inclination i^*, flight-path angle $\gamma_f^* = 0$ (insertion at perigee/apogee, or into a circular orbit), and longitude of ascending node Ω^*. A mission to rendezvous with an orbiting spacecraft would call for the fifth constraint on Ω^*. With i^* and Ω^* specified, the direction of the angular momentum vector of the orbit is fixed, given in an Earth centered inertial (ECI) system by the unit vector $\mathbf{1}_h^{\mathrm{ECI}} = (\sin\Omega^* \sin i^*, -\cos\Omega^* \sin i^*, \cos i^*)^T$. Let $\mathbf{1}_h$ be the representation of $\mathbf{1}_h^{\mathrm{ECI}}$ in the inertial launch plumbline system. The four orbital insertion conditions for a fixed final time are

$$\tfrac{1}{2}\boldsymbol{r}_f^T\boldsymbol{r}_f - \tfrac{1}{2}r_f^{*2} = 0 \tag{78}$$

$$\boldsymbol{r}_f^T\boldsymbol{V}_f = 0 \tag{79}$$

$$\boldsymbol{r}_f^T\mathbf{1}_h = 0 \tag{80}$$

$$\boldsymbol{V}_f^T\mathbf{1}_h = 0 \tag{81}$$

By taking dot products of the transversality condition Eq. (24) with $\boldsymbol{V}_f$, Eq. (25) with $\boldsymbol{V}_f$, then with $\boldsymbol{r}_f$, and making use of the preceding conditions, we again eliminate the multiplier vector $\boldsymbol{\nu}$ and obtain

$$\left(\boldsymbol{V}_f^T\boldsymbol{p}_{r_f}\right)r_f^2 - \left(\boldsymbol{r}_f^T\boldsymbol{p}_{V_f}\right)V_f^2 = 0 \tag{82}$$

$$\boldsymbol{V}_f^T\boldsymbol{p}_{V_f} - V_f^2 = 0 \tag{83}$$

The final time t_f is adjusted the same way by Eq. (76) so that $V_f = V_f^*$.

For staged launch vehicles where all of the stages except for the last one burn to the depletion of the fuel, the burn time of each stage is known for given throttle. The approach in this paper is still applicable with minor modifications. Note that the break points in the finite difference discretization in Sec. III do not have to be equally spaced. In the case of a N-stage vehicle, N different values for the step parameter h may be selected, each for a stage, such that the staging times are located exactly at the last nodes of the stages. The changes of mass and thrust due to staging can then be accommodated easily. The burn time of the last stage is adjusted to meet the final velocity condition.

A number of other orbital insertion conditions were also examined and implemented. However, the principle is the same, thus we will not repeat them in this paper.

B. Open-Loop Solutions

The open-loop solutions are generated using the finite difference (FD) method. For verification and validation of the approach, the trajectories are compared to the ones obtained with other established trajectory optimization tools. The wind velocity $\boldsymbol{V}_w$ is set to be zero in the solutions. The target orbit is a circular orbit at the altitude of 185.2 km (100 n mile). Two orbital inclinations are used: $i^* = 51.6$ deg (the orbital inclination of the International Space Station) and $i^* = 28.5$ deg (the minimum orbital inclination from KSC). A third case is the same circular orbit with $i^* = 51.6$ deg and $\Omega^* = -104$ deg (a five-constraint problem). The initial conditions correspond to those after 5 s of vertical ascent (to clear the tower) with a given launch mass of the X-33. The following path constraints are imposed:

$$|q\alpha| \leq 626.74 \text{ N} \cdot \text{rad/m}^2 \text{ (750 psf} \cdot \text{deg)} \tag{84}$$

$$T \leq 4.0g \tag{85}$$

The engine throttle is set at unity before the thrust acceleration constraint is active. The constraint on dynamic pressure will be added later in closed-loop simulations using the technique presented in

Table 1 Open-loop performance comparison

	FD		SOCS		DIDO	
Trajectory, deg	t_f, s	m_f, kg	t_f, s	m_f, kg	t_f, s	m_f, kg
$i = 51.6$	320.48	38,627.5	321.00	38,476.2	321.12	38,471.8
$i = 28.5$	317.73	39,135.7	318.05	39,011.0	318.18	39,007.2
$\Omega = -104$	322.71	38,220.1	322.33	38,236.5	322.81	38,166.9

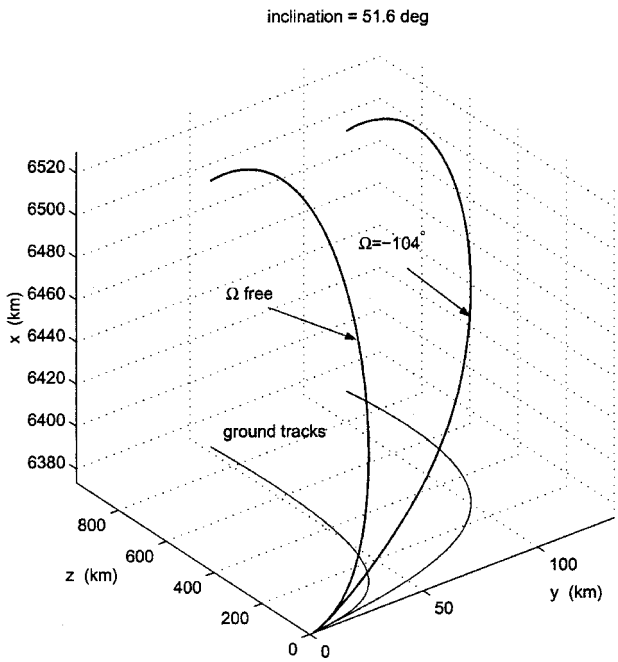

Fig. 2 Three-dimensional ascent trajectories in launch plumbline system.

Sec. II.D. In the following results, 100 nodes were used in FD solutions. For comparison, the solutions for the same missions were also obtained by a collocation software SOCS[14] and by a trajectory optimization software based on a pseudospectral method, DIDO.[15] The SOCS solutions had the same 100 nodes, and the DIDO solutions had 20 nodes. The performance index for both SOCS and DIDO was the final time.

Table 1 summarizes the flight times t_f and final mass m_f of the solutions. The results for all of the missions under different methods are very close. The slight differences among the solutions were mostly due to the discretization errors of the different schemes used in the three methods. This comparison clearly supports the validity of the FD approach.

Figure 2 shows the three-dimensional ascent trajectories and ground tracks for the cases of $i^* = 51.6$ deg and $\Omega = -104$ deg in the inertial launch plumbline coordinate frame. As expected, the trajectory with ascending node constraint had larger out-of-plane motion. Figure 3 shows the angle of attack, pitch angle, and yaw angle along all of the three trajectories by the FD approach. Note that α changes sign a number of times. These sign changes were captured by the criterion stated in Sec. II.C. The constraint on the ascending node usually is one that requires larger yaw and roll maneuvers and, thus, is a demanding constraint. Indeed larger yaw angle is observed in Fig. 3 for that case. The variations of $q\alpha$ and axial thrust acceleration are plotted in Fig. 4. Both path constraints were accurately enforced.

C. Closed-Loop Simulations

In closed-loop simulations, the point-mass vehicle dynamics, atmospheric modeling, propulsive and aerodynamic forces, and winds were simulated by a FORTRAN program. The FD algorithm, which is also implemented in FORTRAN, was called once every second to recalculate the optimal solution based on the current condition. The trajectory simulations used the first data in the optimal body-axes attitude and engine throttle solutions (corresponding to current time) as the guidance commands. No delays or actuator dynamics were simulated. The missions and initial conditions were the same

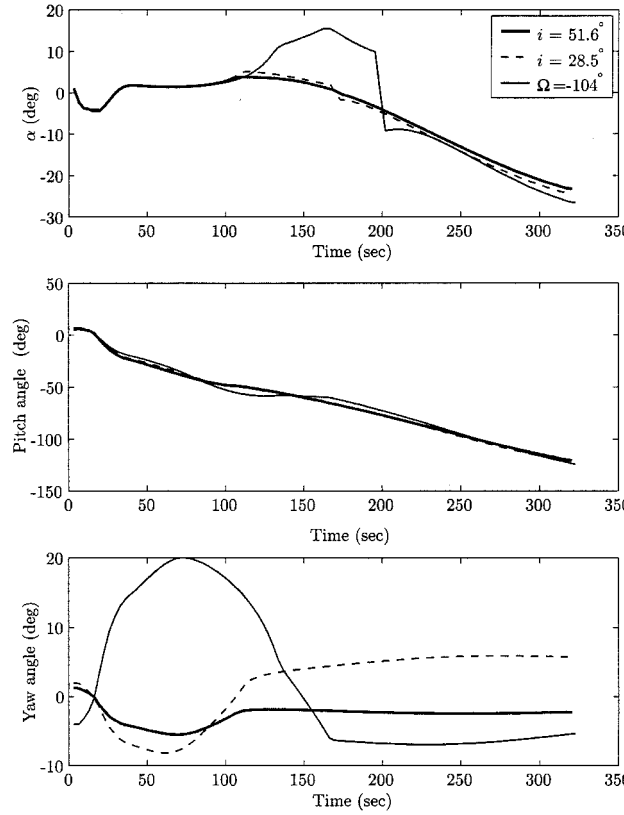

Fig. 3 Angle of attack and pitch and yaw angles in open-loop solutions.

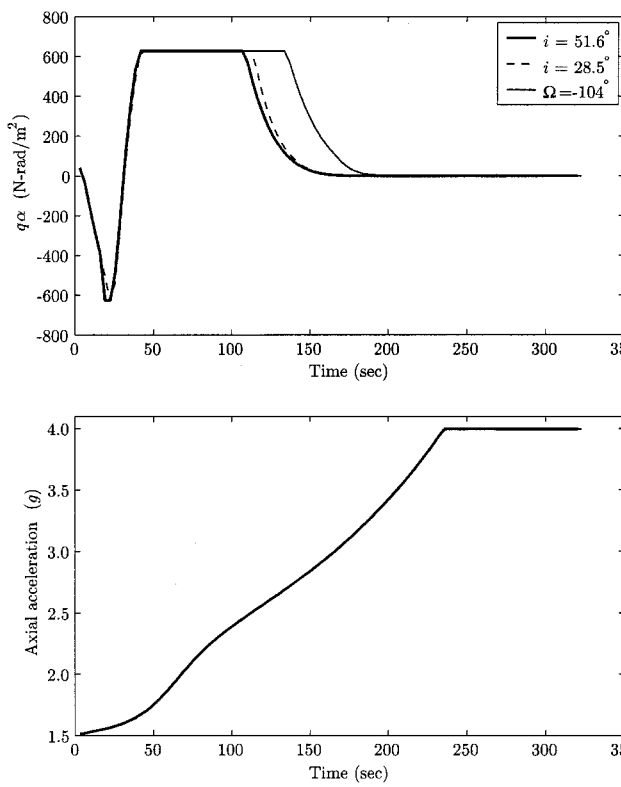

Fig. 4 Variations of $q\alpha$ and axial thrust acceleration in open-loop solutions.

as in the open-loop solutions. In addition to path constraints (84) and (85), a dynamic pressure constraint is also imposed

$$q \leq 18194.4 \text{ N/m}^2 \text{ (380 psf)} \tag{86}$$

Closed-Loop Simulations Without Winds

The previous three missions were repeated in closed-loop simulations with the wind velocity $\boldsymbol{V}_w$ set to be zero. The dynamic pressure

constraint was handled using the technique described in Sec. II.D. There were 30 nodes used in the FD guidance solution. For onboard guidance, fewer nodes can be used than in an open-loop solution because the time to go is decreasing; thus, the accuracy of the solution is increasing while the computation demands would not be unnecessarily high. A converged solution was loaded before "launch" as the starting point. On a desktop computer with a 1-GHz processor, the CPU time that each guidance solution call took ranged from 0.07 to 0.25 s, without any optimization or streamlining of the code. All of the orbital insertion conditions were met accurately. Table 2 lists the flight times and final masses for all three missions. Also in Table 2 are open-loop solutions by SOCS with the addition of the dynamic pressure constraint (86). In all three cases, the closed-loop guided trajectories and the SOCS open-loop solutions have differences in the final mass of about 120 kg or less. These small discrepancies are largely attributed to the differences between closed-loop simulations and open-loop trajectories.

Figure 5 shows the altitude vs inertial velocity along the three trajectories. The variations of dynamic pressure, throttle, and axial thrust acceleration are plotted in Fig. 6. The effectiveness of the feedback approach described in Sec. II.D for enforcing the dynamic pressure constraint is clearly seen: The constraint dynamic pressure constraint (86) is accurately met in all cases, and the feedback throttle modulations are smooth. This is much similar to the "throttle bucket" the shuttle employs for the same purpose, except that the shuttle throttle bucket is preprogrammed (open loop) before launch. In comparison, the SOCS solutions all have considerable zigzags in the throttle when constraint (86) is active (not shown in Fig. 6). From Table 2, the performance (in terms of final mass) of the closed-loop trajectory with the feedback adjustment of the throttle is apparently about the same as that of the open-loop optimal solution. The variations of $q\alpha$ and α are shown in Fig. 7. Figure 8 contains the histories of pitch, yaw, and roll angles along the three trajectories. As expected, the mission with ascending node constraint called for more yaw and roll maneuvers than the other two missions.

Closed-Loop Simulations with Winds

Thus far in the simulations, the wind velocity has been assumed to be zero. To test the ascent guidance algorithm in a more realistic setting, winds were added in the simulations. The wind profiles were based on the measured wind velocities at different altitudes at KSC and then smoothed for guidance purposes. In each simulated trajectory, the same smoothed wind profile was used in both the guidance solution and simulation of the RLV dynamics. We realize that the actual wind the RLV experiences will likely be different from the measured wind because of the time delay between the launch and the time when the measurement was taken. However, the purpose here is to demonstrate how the ascent guidance algorithm would

Table 2 Closed-loop simulations without winds

	FD		SOCS (open-loop)	
Trajectory, deg	t_f, s	m_f, kg	t_f, s	m_f, kg
$i = 51.6$	326.57	38182.4	323.23	38293.6
$i = 28.5$	323.59	38706.7	320.24	38828.6
$\Omega = -104$	327.33	38023.9	324.5	38054.6

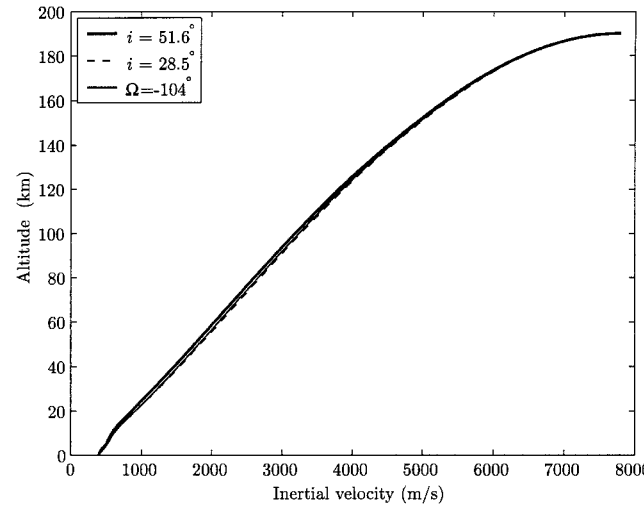

Fig. 5 Altitude vs inertial velocity along closed-loop trajectories.

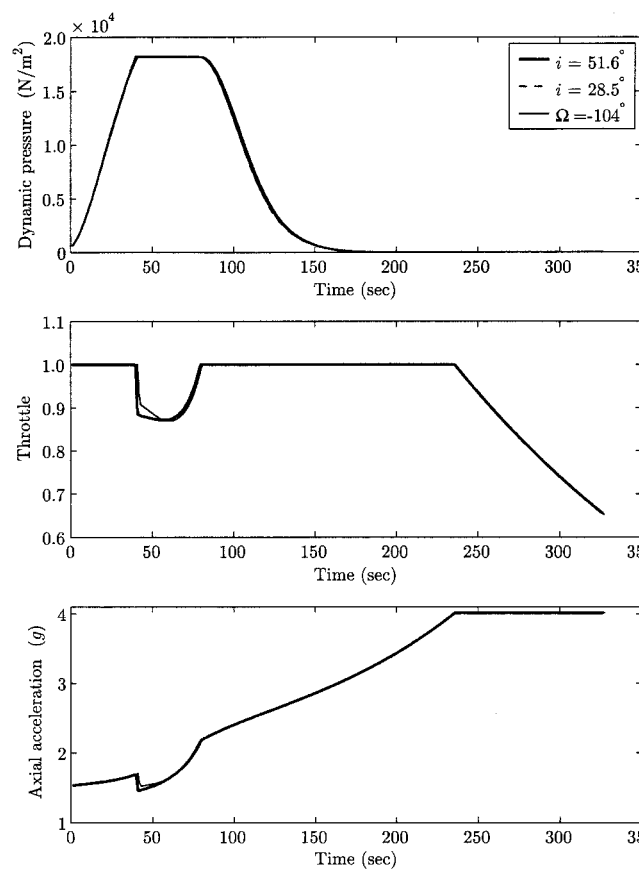

Fig. 6 Dynamic pressure, throttle, and axial acceleration along closed-loop trajectories.

Table 3 Closed-loop simulations with winds (10 wind profiles for each mission)

	$i = 51.6$ deg			$i = 28.5$ deg			$\Omega = -104$ deg		
Parameter	Minimum	Mean	Maximum	Minimum	Mean	Maximum	Minimum	Mean	Maximum
Δr_f, m	−0.0165	0.0128	0.107	−0.0115	0.0152	0.112	−0.0145	0.0484	0.117
ΔV_f, m/s	0.1531	0.2290	0.3007	0.1625	0.2352	0.3327	0.1583	0.1909	0.25388
Δa, km	0.2581	0.3856	0.5063	0.2735	0.3961	0.5602	0.2665	0.3215	0.4275
Δi, deg	$2.4E-5$	$4.6E-5$	$5.9E-5$	$-1.9E-06$	$-4.2E-07$	$5.7E-07$	$-5.8E-4$	$-2.4E-4$	$7.9E-05$
$\Delta\Omega$, deg	N/A	N/A	N/A	N/A	N/A	N/A	$-7.2E-4$	$-2.9E-4$	$9.8E-05$
$\Delta\gamma$, deg	$-3.3E-4$	$1.2E-4$	$1.7E-3$	$-2.6E-4$	$1.3E-4$	$1.5E-3$	$-3.0E-4$	$6.4E-4$	$1.7E-3$
Maximum $\lvert q\alpha\rvert$, N · rad/m^2	919.4	1268.3	1608.3	694.2	1157.8	1641.8	1085.3	1255.8	1690.3
m_f, kg	38010.4	38131.1	38238.5	38507.2	38626.3	38785.4	37789.4	37948.9	38065.3
t_f, s	326.51	328.66	331.21	322.68	326.77	329.51	327.29	329.66	333.19

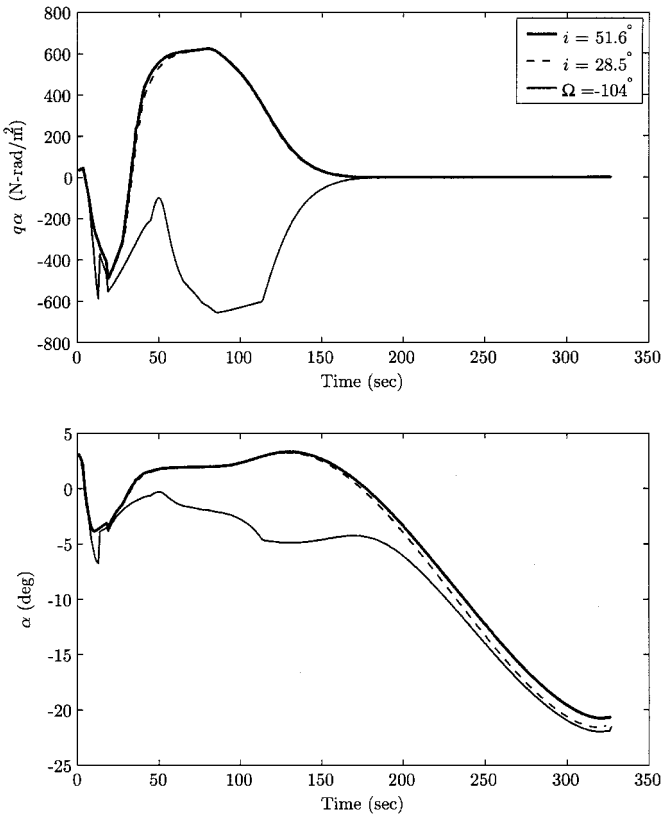

Fig. 7 Variations of $q\alpha$ and α along closed-loop trajectories (no winds).

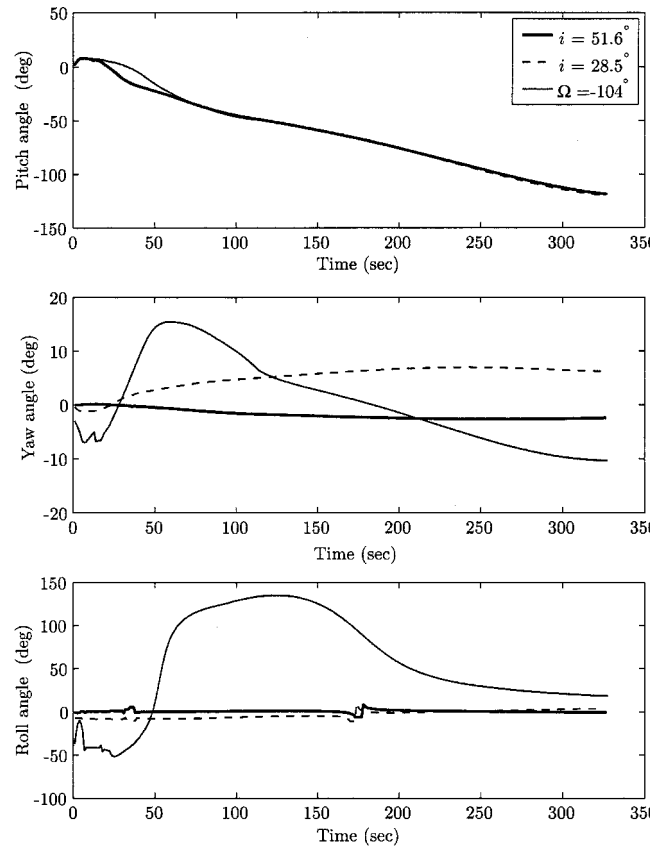

Fig. 8 Variations of Euler angles along closed-loop trajectories (no winds).

perform in the presence of winds should perfect wind information be available. We have also tested the cases when the wind profiles used in guidance algorithm were correlated but not exactly the same as the wind profiles in simulations. In those cases, the enforcement of path constraints is affected by how well the measured wind profiles match the actual wind profiles. With the volatile winds, some of which are quite strong (up to 75 m/s), the limit of $q\alpha$ needs to be higher to fly through the winds. Thus constraint (84) is changed to

$$|q\alpha| \leq 1671.32 \text{ N} \cdot \text{rad/m}^2 \text{ (2000 psf} \cdot \text{deg)} \tag{87}$$

There were 10 trajectories simulated for each of the preceding three missions, with a different wind profile applied in each trajectory. Unlike open-loop ascent guidance, no "prelaunch" adjustments or updates were required for the guidance algorithm for each simulation. The only update for each flight was the wind profile. The closed-loop ascent guidance algorithm automatically incorporates the wind data in the guidance solution. The starting point was still the same zero-wind solution used before. The CPU time for each guidance cycle did not differ from the earlier recorded value by any noticeable margin.

The simulation results are summarized in Table 3. The first six quantities with Δ are the orbital insertion condition errors. The minimum, average, and maximum values of these errors among the 10 trajectories for each mission are listed. The altitude errors were all within 0.2 m and velocity errors within 0.4 m/s. The quantity Δa is the error in semimajor axis of the final orbit, which is a parameter combining the effects of Δr_f and ΔV_f. The maximum of Δa was no greater than 0.6 km. All other errors were very small as well. The peak values of $|q\alpha|$ were essentially all within the specified range. The product of $q\beta$, where β is the sideslip angle, was not listed because it remained practically zero because of the roll maneuvers by the RLV to "fly into the wind." Even when the wind profiles used in simulations were different from but correlated to the wind profiles used in the guidance solution, such as in actual launch, the peak values of $|q\beta|$ in general were still significantly smaller than those of $|q\alpha|$ (not shown here). The average final masses for all three missions were quite close to those in Table 2 for the trajectories without winds. Note that whether or not the wind data are included in the optimal guidance solutions can make a sizable difference in the performance. If the wind profiles are merely used outside the guidance solution to limit the attitude guidance commands for observing the $q\alpha$-constraint, the results using the X-33 vehicle model can mean up to 500-kg less payload delivered into the same orbit.

Tests on other ascent missions and abort-to-orbit missions, not reported in this paper, have also been performed. The test results were all similar to those described earlier.

V. Conclusions

A comprehensive treatment of the problem of optimal ascent of launch vehicles through the atmosphere is provided. Special attention has been given to the issues that impact on viability, complexity, and reliability in onboard implementation. The classical FD method is shown to be well suited for fast solution of the constrained optimal three-dimensional ascent problem. The convergence speed is such that closed-loop endoatmospheric ascent guidance should be achievable within the capability of the next-generation flight computers. A series of open-loop and closed-loop tests using the data of a reusable launch vehicle was designed to demonstrate the performance of the algorithm. The results clearly establish the feasibility of closed-loop endoatmospheric ascent guidance of rocket-powered launch vehicles.

Appendix A: Costate Equations for Three-Dimensional Optimal Ascent

The notation used hereinafter is the same as defined in Secs. II.A and II.B. Let $\rho_r = \partial\rho/\partial r$, $T_r = \partial T/\partial r$, $C_\rho = \rho_0 \rho S_{\text{ref}} R_0/m(t)$, and

$$A_{\rho r} = \frac{V_r^2 S_{\text{ref}} R_0 \rho_0 C_A \rho_r}{2m(t)}, \qquad N_{\rho r} = \frac{V_r^2 S_{\text{ref}} R_0 \rho_0 C_N \rho_r}{2m(t)}$$

$$C_{A_\alpha} = \frac{\partial C_A}{\partial \alpha}, \qquad C_{N_\alpha} = \frac{\partial C_N}{\partial \alpha}, \qquad C_{A_{\text{Mach}}} = \frac{\partial C_A}{\partial \text{Mach}}$$

$$C_{N_{\text{Mach}}} = \frac{\partial C_N}{\partial \text{Mach}}, \qquad a_{pvb} = \boldsymbol{p}_V^T \mathbf{1}_b = p_V \cos(\Phi - \alpha)$$

$$a_{pvn} = \boldsymbol{p}_V^T \mathbf{1}_n = p_V \sin(\Phi - \alpha)$$

Denote the altitude-dependent speed of sound by $V_s(r)$. The complete costate equations (21) and (22), after much vector differentiation and simplification, are given by

$$\boldsymbol{p}_r' = \frac{1}{r^3}\boldsymbol{p}_V - \left[\frac{3a_{pvb}}{r^4} - a_{pvn}\left(T_r - A_{\rho r} + \frac{1}{2V_r}C_\rho V_s^2 C_{A_{\text{Mach}}}\frac{\partial V_s}{\partial r}\right) + a_{pvn}\left(N_{\rho r} - \frac{1}{2V_r}C_\rho V_s^2 C_{N_{\text{Mach}}}\frac{\partial V_s}{\partial r}\right)\right]\frac{\boldsymbol{r}}{r} + C_\rho\bar{\omega}_E \times \left\{a_{pvb}\left[\left(C_A + \frac{1}{2V_r}V_s C_{A_{\text{Mach}}}\right)\boldsymbol{V}_r + \frac{1}{2}C_{A_\alpha}V_r^2\frac{\partial\alpha}{\partial\boldsymbol{V}}\right] - a_{pvn}\left[\left(C_N + \frac{1}{2V_r}V_s C_{N_{\text{Mach}}}\right)\boldsymbol{V}_r + \frac{1}{2}C_{N_\alpha}V_r^2\frac{\partial\alpha}{\partial\boldsymbol{V}}\right]\right\} \tag{A1}$$

$$\boldsymbol{p}_V' = -\boldsymbol{p}_r + C_\rho\left[a_{pvb}\left(C_A + \frac{1}{2V_r}V_s C_{A_{\text{Mach}}}\right) - a_{pvn}\left(C_N + \frac{1}{2V_r}V_s C_{N_{\text{Mach}}}\right) + \frac{1}{2}\left(a_{pvb}C_{A_\alpha} - a_{pvn}C_{N_\alpha}\right)\frac{\cos\alpha}{\sin\alpha}\right] \times \boldsymbol{V}_r - \frac{C_\rho V_r}{2\sin\alpha}\left(a_{pvb}C_{A_\alpha} - a_{pvn}C_{N_\alpha}\right)\mathbf{1}_b \tag{A2}$$

where, from the first equation in Eq. (12) and $\boldsymbol{V}_r = \boldsymbol{V} - \bar{\omega}_E \times \boldsymbol{r} - \boldsymbol{V}_w$,

$$\frac{\partial\alpha}{\partial\boldsymbol{V}} = \frac{1}{V_r\sin\alpha}\left(\cos\alpha\mathbf{1}_{V_r} - \mathbf{1}_b\right) \tag{A3}$$

Note that the costate equation (A1) has been simplified by recognizing that

$$\left(\mathbf{1}_b\mathbf{1}_b^T + \mathbf{1}_n\mathbf{1}_n^T - I_3\right)\boldsymbol{p}_V = 0 \tag{A4}$$

because $\mathbf{1}_b$, $\mathbf{1}_n$, and $\boldsymbol{p}_V$ are in the same plane in the optimal solution and $\mathbf{1}_b$ and $\mathbf{1}_n$ are unit vectors that are orthogonal to each other.

Appendix B: Analytical Vacuum Optimal Ascent Solutions

The analytical vacuum optimal ascent guidance solution combines a number of elegant results in optimal vacuum trajectory studies over the past three decades, as summarized in Ref. 6. The key ingredients are the linear gravity approximation, the closed-form solution of the costate equation, and closed-form solution of the state equation using quadrature formulas. At the beginning of each guidance update cycle, let $\boldsymbol{r}_0$ be the dimensional position vector. The gravity acceleration $\boldsymbol{g}$ from this point on is approximated by a linear function of the position vector $\boldsymbol{r}$,

$$\boldsymbol{g} = -\left(\mu_E/r_0^3\right)\boldsymbol{r} = -\omega^2\boldsymbol{r} \tag{B1}$$

where μ_E is the gravitational parameter of the Earth and ω is the Schuler frequency at r_0. In an ascent guidance problem, the correct direction of the gravity is more important than the accuracy of its magnitude. This approximation preserves the change of direction of the gravitational acceleration with $\boldsymbol{r}$. The magnitude of $\boldsymbol{g}$ will be slightly different from that of a Newtonian central gravity field. However, when r_0 is continuously updated by the radius at beginning of each guidance cycle, the effect of this difference will be negligible.

Let $g_0 = \mu_E/r_0^2$ be the magnitude of the gravity acceleration at r_0. We normalize the equations of motion (1) and (2) with unit distance r_0, unit time $\sqrt{(r_0/g_0)}$, and unit velocity $\sqrt{(r_0 g_0)}$. The dimensionless equations of motion, with $\boldsymbol{A} = \boldsymbol{N} = 0$ for vacuum flight, become

$$\dot{\boldsymbol{r}} = \boldsymbol{V} \tag{B2}$$

$$\dot{\boldsymbol{V}} = -\boldsymbol{r} + T(\tau)\mathbf{1}_b \tag{B3}$$

where the Schuler frequency ω has become unity in the normalized time and $T(\tau) = T_{\text{vac}}/m(\tau)g_0$ with τ as the normalized time. Note that this normalization is done in each guidance cycle with the r_0 being the radius at the beginning of that cycle. The Hamiltonian now is

$$H = \boldsymbol{p}_r^T\boldsymbol{V} + \boldsymbol{p}_V^T[-\boldsymbol{r} + T(\tau)\mathbf{1}_b] + \mu\left(\mathbf{1}_b^T\mathbf{1}_b - 1\right) \tag{B4}$$

The optimality condition from $\partial H/\partial\mathbf{1}_b = 0$ yields

$$\mathbf{1}_b^* = -[T(\tau)/2\mu]\boldsymbol{p}_V \tag{B5}$$

The sufficient condition for the optimality condition (23) in this case is $\partial^2 H/\partial\mathbf{1}_b^2 = 2\mu I_3 < 0$. We have $\mu < 0$, hence, the well-known result that the optimal thrust direction in Eq. (B5) must be aligned with that of the primer vector $\boldsymbol{p}_V$ (Ref. 12). The costate equations (21) and (22) now become

$$\dot{\boldsymbol{p}}_r = \boldsymbol{p}_V \tag{B6}$$

$$\dot{\boldsymbol{p}}_V = -\boldsymbol{p}_r \tag{B7}$$

The costate equations have closed-form solution of

$$\begin{bmatrix}\boldsymbol{p}_V(\tau)\\ -\boldsymbol{p}_r(\tau)\end{bmatrix} = \begin{bmatrix}\cos\tau I_3 & \sin\tau I_3\\ -\sin\tau I_3 & \cos\tau I_3\end{bmatrix}\begin{bmatrix}\boldsymbol{p}_{V_0}\\ -\boldsymbol{p}_{r_0}\end{bmatrix} \triangleq \Omega(\tau)\begin{bmatrix}\boldsymbol{p}_{V_0}\\ -\boldsymbol{p}_{r_0}\end{bmatrix} \tag{B8}$$

where $\boldsymbol{p}_{V_0}$ and $\boldsymbol{p}_{r_0}$ are the (unknown) initial conditions for the costate. Define

$$\boldsymbol{I}_c(\tau) = \int_0^\tau \mathbf{1}_{pV}(\zeta)\cos\zeta\, T(\zeta)\,\mathrm{d}\zeta \triangleq \int_0^\tau \dot{\boldsymbol{I}}_c(\zeta)\,\mathrm{d}\zeta \tag{B9}$$

$$\boldsymbol{I}_s\tau) = \int_0^\tau \mathbf{1}_{pV}(\zeta)\sin\zeta\, T(\zeta)\,\mathrm{d}\zeta \triangleq \int_0^\tau \dot{\boldsymbol{I}}_s(\zeta)\,\mathrm{d}\zeta \tag{B10}$$

Note that thrust acceleration $T(\cdot)$ is time varying because the mass is changing. Also the engine throttle may be adjusted according to Eq. (45) for enforcing the axial acceleration constraint. It can be easily verified that the state equations have the solution of[20]

$$\begin{bmatrix}\boldsymbol{r}(\tau)\\ \boldsymbol{V}(\tau)\end{bmatrix} = \Omega(\tau)\begin{bmatrix}\boldsymbol{r}_0\\ \boldsymbol{V}_0\end{bmatrix} + \Gamma(\tau)\begin{bmatrix}\boldsymbol{I}_c(\tau)\\ \boldsymbol{I}_s(\tau)\end{bmatrix} \tag{B11}$$

where

$$\Gamma(\tau) = \begin{bmatrix}\sin\tau I_3 & -\cos\tau I_3\\ \cos\tau I_3 & \sin\tau I_3\end{bmatrix} \tag{B12}$$

The integrals $\boldsymbol{I}_c$ and $\boldsymbol{I}_s$ can be evaluated by a numerical quadrature scheme. Calise et al.[6] use Simpson's rule. We opt to use Milne's rule because it only increases computation by a small margin, but offers considerably higher precision than Simpson's rule (see Ref. 18). Let $\delta = \tau_{\text{togo}}/4$ where τ_{togo} is the dimensionless time to go till main engine cutoff. The value of τ_{togo} is known for a fixed final-time problem or to be determined in a problem with free final time. With the Milne's rule, we have

$$\boldsymbol{I}_i(\tau_{\text{togo}}) = (\tau_{\text{togo}}/90)[7\dot{\boldsymbol{I}}_i(0) + 32\dot{\boldsymbol{I}}_i(\delta) + 12\dot{\boldsymbol{I}}_i(2\delta) + 32\dot{\boldsymbol{I}}_i(3\delta) + 7\dot{\boldsymbol{I}}_i(4\delta)], \qquad i = c, s \tag{B13}$$

Milne's rule has a truncation error proportional to δ^7, whereas for the same τ_{togo} Simpson's rule has a truncation error proportional to $(2\delta)^5 = 32\delta^5$ for the same δ. With the thrust integrals given in Eq. (B13) and the costate given in solution (B8) as function of its initial conditions, the state is found in closed form from Eq. (B11). Therefore, the final state and costate are explicitly functions of $\boldsymbol{p}_{V_0}$ and $\boldsymbol{p}_{r_0}$. Consequently, the total six terminal conditions (14) and (27) are functions of the six unknowns $\boldsymbol{p}_{V_0}$ and $\boldsymbol{p}_{r_0}$. In a problem with free final time, the seventh unknown is τ_{togo}, and the seventh condition is from condition (26). It can be shown that in a minimum-time problem for Keplerian orbit insertion, condition (26) is automatically satisfied.[21] Such a discussion has been conspicuously missing in previous literature on optimal vacuum guidance. Note that this conclusion is not necessary true for other performance indices or terminal conditions. In those cases, condition (26) needs to be examined and enforced if necessary.

For minimum-time problem where $J = \phi = t_f$, the costate can be scaled by an arbitrary positive constant without changing any necessary conditions for the optimal control problem. Therefore, $\boldsymbol{p}_{V_0}$ and $\boldsymbol{p}_{r_0}$ are not completely independent because one can always scale the costate, for instance, so that $\|\boldsymbol{p}_0\| = \|(\boldsymbol{p}_{V_0}^T \boldsymbol{p}_{r_0}^T)^T\| = 1$. One of the six components of $\boldsymbol{p}_0$ can be determined by the rest. However, this still leaves us the ambiguity of determining the sign of that component of $\boldsymbol{p}_0$. We avoid this problem by still treating the problem as a seven-unknown problem ($\boldsymbol{p}_0$ plus τ_{togo}) and adding a trivial condition to make up for the seventh terminal condition

$$\|\boldsymbol{p}(\tau_f)\| = 1 \tag{B14}$$

From the costate equations (B6) and (B7), it can be shown by simple differentiation that

$$\frac{\mathrm{d}\|\boldsymbol{p}(\tau)\|}{\mathrm{d}\tau} = 0 \tag{B15}$$

Therefore, $\|\boldsymbol{p}(\tau)\| =$ constant. Condition (B14), thus, will always be trivially satisfied if we scale $\boldsymbol{p}_0$ to have $\|\boldsymbol{p}_0\| = 1$ (and we must, for condition (B14) to be met).

In summary, the minimum-time vacuum ascent guidance problem becomes a root-finding problem with seven unknowns ($\boldsymbol{p}_0$ and τ_{togo}), six constraints (14) and (27) plus one "easy" constraint (B14). Through the use of quadrature (B13), all of the final state $\boldsymbol{x}_f$ and costate $\boldsymbol{p}_f$ are explicit functions of the seven unknowns. The modified Newton method again works very well, and the convergence occurs rapidly with almost any initial guesses that do not result in totally wrong initial thrust direction. All of the Jacobians needed in the Newton iterations can be obtained analytically in a similar fashion as in Ref. 6. Once $\boldsymbol{p}_0$ and τ_{togo} are found, the entire state and costate histories are determined, and the current commanded thrust direction is in the direction of $\boldsymbol{p}_{V_0}$.

Acknowledgment

This research has been supported in part by the NASA Contract NAS8-01105 under the Space Launch Initiative Program. The Technical Officer is Jack Mulqueen.

References

[1]Smith, I. E., "General Formulation of the Iterative Guidance Mode," NASA TM X-53414, March 1966.

[2]McHenry, R. L., Brand, T. J., Long, A. D., Cockrell, B. F., and Thibodeau, J. R., III, "Space Shuttle Ascent Guidance, Navigation, and Control," *Journal of the Astronautical Sciences*, Vol. 27, No. 1, 1979, pp. 1–38.

[3]Hanson, J. M., Shrader, M. W., and Cruzen, A., "Ascent Guidance Comparisons," *Proceedings of the AIAA Guidance, Navigation, and Control Conference*, Vol. 1, AIAA, Washington, DC, 1994, pp. 230–240.

[4]Brown, K. R., and Johnson, G. W., "Real-Time Optimal Guidance," *IEEE Transactions on Automatic Control*, Vol. AC-12, No. 5, 1967, pp. 501–506.

[5]Leung, M. S. K., and Calise, A. J., "Hybrid Approach to Near-Optimal Launch Vehicle Guidance," *Journal of Guidance, Control, and Dynamics*, Vol. 17, No. 5, 1994, pp. 881–888.

[6]Calise, A. J., Melamed, N., and Lee, S., "Design and Evaluation of a Three-Dimensional Optimal Ascent Guidance Algorithm," *Journal of Guidance, Control, and Dynamics*, Vol. 21, No. 6, 1998, pp. 867–875.

[7]Gath, P. F., and Calise, A. J., "Optimization of Launch Vehicle Ascent Trajectories with Path Constraints and Coast Arcs," *Journal of Guidance, Control, and Dynamics*, Vol. 24, No. 2, 2001, pp. 296–304.

[8]Gath, P. F., and Well, K. H., "HISTOS Technical Report 1," Inst. of Flight Mechanics and Control, Univ. of Stuttgart, Stuttgart, Germany, June 1999.

[9]Calise, A. J., and Brandt, N., "Generation of Launch Vehicle Abort Trajectories Using a Hybrid Optimization Method," AIAA Paper 2002-4560, Aug. 2002.

[10]Bryson, A. E., and Ho, Y. C., *Applied Optimal Control*, Hemisphere, Washington, DC, 1975, pp. 42–87.

[11]Vinh, N. X., "General Theory of Optimal Trajectory for Rocket Flight in a Resisting Medium," *Journal of Optimization Theory and Applications*, Vol. 11, No. 2, 1973, pp. 189–202.

[12]Lawden, D. F., *Optimal Trajectories for Space Navigation*, Butterworth, London, 1963, pp. 54–68.

[13]Lu, P., "Nonlinear Systems with Control and State Constraints," *Optimal Control Applications and Methods*, Vol. 18, 1997, pp. 313–326.

[14]Betts, J. T., and Huffman, W. P., Sparse Optimal Control Software, Ver. 5.0, The Boeing Co., Seattle, WA, June 2000.

[15]Ross, I. M., and Fahroo, F., "User's Manual for DIDO 2001 (α): A MATLAB^{TM} Application Package for Dynamic Optimization," Naval Postgraduate School, TR NPS-AA-01-003, Monterey, CA, Oct. 2001.

[16]Corvin, M. A., "Ascent Guidance for a Winged Boost Vehicle," NASA CR-172083, Aug. 1988.

[17]Keller, H. B., *Numerical Methods for Two-Point Boundary-Value Problems*, Blaisdell, Waltham, MA, 1968, pp. 91–100.

[18]Stoer, J., and Bulirsch, R., *Introduction to Numerical Analysis*, Springer-Verlag, New York, 1980, pp. 256–262.

[19]Press, W. H., Teukolsky, S. A., Vetterling, W. T., and Flannery, B. P., *Numerical Recipes in FORTRAN*, Cambridge Univ. Press, Cambridge, England, U.K., 1992, pp. 753–759.

[20]McAdoo, S. F., Jezewski, D. J., and Dawkins, G. S., "Development of a Method for Optimal Maneuver Analysis of Complex Space Missions," NASA TN D-7882, April 1975.

[21]Dukeman, G. A., "Atmospheric Ascent Guidance for Rocket-Powered Launch Vehicles," AIAA Paper 2002-4559, Aug. 2002.