JOURNAL OF GUIDANCE, CONTROL, AND DYNAMICS
Vol. 24, No. 2, March–April 2001

Costate Estimation by a Legendre Pseudospectral Method

Fariba Fahroo* and I. Michael Ross†
Naval Postgraduate School, Monterey, California 93943

We present a Legendre pseudospectral method for directly estimating the costates of the Bolza problem encountered in optimal control theory. The method is based on calculating the state and control variables at the Legendre–Gauss–Lobatto (LGL) points. An *N*th degree Lagrange polynomial approximation of these variables allows a conversion of the optimal control problem into a standard nonlinear programming (NLP) problem with the state and control values at the LGL points as optimization parameters. By applying the Karush–Kuhn–Tucker (KKT) theorem to the NLP problem, we show that the KKT multipliers satisfy a discrete analog of the costate dynamics including the transversality conditions. Indeed, we prove that the costates at the LGL points are equal to the KKT multipliers divided by the LGL weights. Hence, the direct solution by this method also automatically yields the costates by way of the Lagrange multipliers that can be extracted from an NLP solver. One important advantage of this technique is that it allows a very simple way to check the optimality of the direct solution. Numerical examples are included to demonstrate the method.

I. Introduction

MANY different computational methods exist for solving optimal control problems,[1,2] but they can be grouped into two major categories: indirect and direct methods. In indirect methods, the necessary optimality conditions derived from the minimum principle are solved to obtain the optimal trajectory. These methods are not typically used to solve complex problems due to 1) their inherent small radii of convergence and 2) the additional labor required in deriving the optimality conditions. Direct methods can be basically described as solving the optimal control problem by discretizing it to a parameter optimization problem and then solving the resulting nonlinear programming problem (NLP). The conversion to a parameter optimization problem can be classified into two major categories: 1) parameterization of the control variable only and 2) parameterization of both control and state variables. This paper deals with the latter category. Also note that if the controls can be eliminated by differential inclusion only the state variables need to be discretized.[3,4] In most direct methods,[2] the conversion to a parameter optimization problem is achieved by first dividing the time interval into a prescribed number of subintervals whose endpoints are called nodes. The unknowns are the value of the controls and the states at these nodes, the state and control parameters. The cost function and the state equations can be expressed in terms of these parameters, which effectively reduce the optimal control problem to an NLP that can be solved by a standard nonlinear programming code. The time histories of both the control and the state variables can be obtained by using an interpolation scheme. In most collocation schemes, linear or cubic splines are used as the interpolating polynomials.[5,6] To impose the state differential equations, some form of implicit integration scheme is used, among which a popular one is the Hermite–Simpson scheme (Ref. 5). Gauss–Lobatto quadrature rules such as trapezoidal, Simpson's, or higher-order rules with Jacobi polynomials as the interpolant are also used for collocation.[7]

Instead of using piecewise-continuous polynomials as the interpolant between prescribed subintervals, global orthogonal polynomials such as Legendre and Chebyshev polynomials can be used for the approximation of the control and state variables.[4,8–10] These polynomials are used extensively in spectral methods for solving fluid dynamics problems,[11,12] but their use in solving optimal control problems has created a new way of transforming them to NLP problems. One particular merit of the use of orthogonal polynomials is their close relationship to Gauss-type integration rules. This relationship can be used to derive simple rules for transforming the original optimal control problem to a system of algebraic equations. The efficiency and simplicity of these rules are best demonstrated in the spectral collocation method that Elnagar et al.,[13] Elnagar and Razzaghi,[14] and Fahroo and Ross[4,15] have recently employed to solve a general class of optimal control problems. In this method, polynomial approximations of the state and control variables are considered where Lagrange polynomials are the trial functions and the unknown coefficients are the values of the state and control variables at the Legendre-Gauss-Lobatto (LGL) points. It is well known that this choice of collocation points yield superior results for interpolation of functions to the ones obtained from equidistant points.[16] By using the properties of the Lagrange polynomials, the state equations and the control constraints are readily transformed to algebraic equations. The state differential constraints are imposed by evaluating the functions at the LGL points and using a differentiation matrix that is obtained by taking the analytic derivative of the interpolating polynomials and evaluating them at the LGL points. In this sense, this method of imposing the state equations is in marked contrast to the numerical integration techniques that are used to approximate the differential equations in other collocation schemes. The integral cost function can also be discretized by Gauss–Lobatto quadrature rules, which provide highly accurate results for approximating integrals.[16] Therefore, this method unifies the process of discretization of the differential equations and the integral cost functions. In other words, the Bolza problem need not be converted to a Mayer problem, a technique typically employed in the traditional collocation schemes.

The use of the differentiation matrix, which is a central idea in pseudospectral methods, has provided us two important extensions[4,17] of these methods: conversion of the differential state dynamics to differential inclusions and indirect collocation methods for solving for states and costate variables simultaneously. The traditional implementation of differential inclusions[3] relies on a low-order explicit Euler-type approximation, but the differentiation matrix in spectral collocation can easily be used in conversion of the state equations to differential inclusion without sacrificing accuracy or increasing the number of NLP variables.[4] In the second extension,[17] the state and the costate variables are approximated by a spectral collocation method, and the two-point boundary value problem is posed as an NLP, which can then be readily solved.

In this paper, we investigate the relationship between the costate variables and the Lagrange multipliers used in the discrete method. By applying the Karush–Kuhn–Tucker (KKT) theorem, we prove

Received 23 October 1998; revision received 22 May 2000; accepted for publication 25 May 2000.

*Assistant Professor, Department of Mathematics, Code Ma/Ff; ffahroo@nps.navy.mil. Senior Member AIAA.

†Associate Professor, Department of Aeronautics and Astronautics, Code AA/Ro; imross@nps.navy.mil. Associate Fellow AIAA.

that the KKT multipliers satisfy the same conditions as those obtained by collocating the costate equations. This key result does not hold for all collocation methods. For example, the transcription method using Hermite interpolation and Simpson quadrature rule yields a discrete adjoint system which is of lower order of accuracy than the state approximations.[18] In contrast, our discretization method offers the same order of accuracy for the states and costates. As a result, the costates can be determined quite accurately at the LGL points by simply dividing the KKT multipliers by the LGL weights.

In this sense, the present method for costate estimations is significantly simpler, both in presentation and implementation, than previous methods.[6,18–20] One important advantage of use of this spectral method is in the ease and efficiency of estimating the costates from the results of the direct method. Because convergence to a solution in indirect methods is highly dependent on the initial guess, using the direct method for estimating the costates can prove quite valuable particularly because in most cases a good guess to the costates is not available. For most problems, direct methods can be readily obtained with a wider radius of convergence; however, the results may not be as accurate as the ones obtained from indirect methods. If one can find an approximation to the costates from the direct results, then this combination of direct and indirect method can lead to a more accurate and efficient way of solving optimal control problems.[6]

II. Problem Formulation

Consider the following optimal control problem. Determine the control function, $\boldsymbol{u}(\tau)$, and the corresponding state trajectory, $\boldsymbol{x}(\tau)$, that minimize the Bolza cost function:

$$\mathcal{J}(\boldsymbol{u}, \boldsymbol{x}, \tau_f) = \mathcal{M}[\boldsymbol{x}(\tau_f), \tau_f] + \int_{\tau_0}^{\tau_f} \mathcal{L}[\boldsymbol{x}(\tau), \boldsymbol{u}(\tau)]\,\mathrm{d}\tau \tag{1}$$

where $\boldsymbol{x} \in R^n$ and $\boldsymbol{u} \in R^m$ are subject to the state dynamics

$$\dot{\boldsymbol{x}}(\tau) = \boldsymbol{f}[\boldsymbol{x}(\tau), \boldsymbol{u}(\tau)], \qquad \tau \in [\tau_0, \tau_f] \tag{2}$$

and boundary conditions

$$\psi_0[\boldsymbol{x}(\tau_0), \tau_0] = \boldsymbol{0} \tag{3}$$

$$\psi_f[\boldsymbol{x}(\tau_f), \tau_f] = \boldsymbol{0} \tag{4}$$

where $\psi_0 \in R^p$ with $p \le n$ and $\psi_f \in R^q$ with $q \le n$. We consider an autonomous system because an extension to a nonautonomous system is straightforward. For simplicity, we do not consider state constraints, although the direct method can handle them easily.[4,10,13,15]

Possible control inequality constraints are formulated as

$$\boldsymbol{g}[\boldsymbol{u}(\tau)] \le \boldsymbol{0}, \qquad \boldsymbol{g} \in R^r \tag{5}$$

where $\partial\boldsymbol{g}/\partial\boldsymbol{u}$ has full rank.

The Lagrange multiplier theory for this problem allows us to adjoin the state equation and constraints to the cost function by the time-dependent multipliers $\boldsymbol{\lambda}(\tau)$ and constant multipliers $\boldsymbol{\nu}_0 \in R^p$ and $\boldsymbol{\nu}_f \in R^q$ to form the augmented cost function:

$$\bar{\mathcal{J}} = \mathcal{M}[\boldsymbol{x}(\tau_f), \tau_f] + \boldsymbol{\nu}_0^T\psi_0 + \boldsymbol{\nu}_f^T\psi_f + \int_{\tau_0}^{\tau_f} [\mathcal{L}(\boldsymbol{x}, \boldsymbol{u}) + \boldsymbol{\lambda}^T(\tau)\{f[\boldsymbol{x}(\tau), \boldsymbol{u}(\tau)] - \dot{\boldsymbol{x}}\} + \boldsymbol{\mu}^T(\tau)\boldsymbol{g}(\tau)]\,\mathrm{d}\tau \tag{6}$$

In terms of the augmented Hamiltonian defined as

$$\mathcal{H}(\boldsymbol{x}, \boldsymbol{\lambda}, \boldsymbol{u}) = \boldsymbol{\lambda}^T\boldsymbol{f} + \mathcal{L} + \boldsymbol{\mu}^T\boldsymbol{g} \tag{7}$$

the necessary optimality conditions are given by

$$\frac{\partial\mathcal{H}}{\partial\boldsymbol{u}} = \boldsymbol{0}, \qquad \boldsymbol{\mu}^T\boldsymbol{g} = 0, \qquad \boldsymbol{\mu} \ge \boldsymbol{0} \tag{8}$$

where $\boldsymbol{\lambda}(t)$ are governed by the costate dynamics and the transversality conditions

$$\dot{\boldsymbol{\lambda}} = -\frac{\partial\mathcal{H}}{\partial\boldsymbol{x}} \tag{9}$$

$$\boldsymbol{\lambda}(\tau_0) = -\left[\frac{\partial\psi_0}{\partial\boldsymbol{x}(\tau_0)}\right]^T\boldsymbol{\nu}_0 \tag{10}$$

$$\boldsymbol{\lambda}(\tau_f) = \frac{\partial\mathcal{M}}{\partial\boldsymbol{x}(\tau_f)} + \left[\frac{\partial\psi_f}{\partial\boldsymbol{x}(\tau_f)}\right]^T\boldsymbol{\nu}_f \tag{11}$$

$$\mathcal{H}(\tau_f) = -\left[\frac{\partial\mathcal{M}}{\partial\tau_f} + \boldsymbol{\nu}_f^T\frac{\partial\psi_f}{\partial\tau_f}\right] \tag{12}$$

III. Legendre Pseudospectral Method

In this section, we present a Legendre pseudospectral method (Legendre spectral collocation method)[11–15] for solving the optimal control problem formulated in the preceding section. The basic idea of this method consists of two steps: First, we seek global polynomial approximations for the state and control functions in terms of their values at the LGL points. Second, we find equations that these approximations satisfy. To obtain these equations, we impose the condition that the state equations are satisfied exactly by these approximations at the LGL points. Unlike the collocation methods that use piecewise-continuous functions on the arbitrary subintervals (Hermite–Simpson scheme, for example), the polynomials (the trial functions) used in this method are globally interpolating Lagrange polynomials obtained from the orthogonal Legendre polynomials. Therefore, by definition, the coefficients of the polynomial expansion are exactly the values of the functions at the LGL points. By the use of the same LGL points, the integral and differential portions of the problem are discretized: The cost function integral is calculated by the LGL quadrature integration rule, whereas the time derivative of the approximate state vector $\dot{\boldsymbol{x}}^N(\tau)$ is expressed in terms of the approximate state vector $\boldsymbol{x}^N(\tau)$ at the collocation points by the use of a differentiation matrix. In this manner, the optimal control problem is transformed to an NLP problem for the values of the states and the controls at the LGL nodes.

Because the problem presented in the preceding section is formulated over the time interval $[\tau_0, \tau_f]$, and the LGL points lie in the interval $[-1, 1]$, we use the following transformation to express the problem for $t \in [t_0, t_N] = [-1, 1]$:

$$\tau = [(\tau_f - \tau_0)t + (\tau_f + \tau_0)]/2 \tag{13}$$

The use of the symbol t_N (which maps τ_f) will be apparent shortly. It follows that by using Eq. (13), Eqs. (1–5) can be replaced by

$$\mathcal{J}[\boldsymbol{x}(\cdot), \boldsymbol{u}(\cdot), \tau_f] = \mathcal{M}[\boldsymbol{x}(1), \tau_f] + \frac{\tau_f - \tau_0}{2}\int_{-1}^{1} \mathcal{L}[\boldsymbol{x}(t), \boldsymbol{u}(t)]\,\mathrm{d}t \tag{14}$$

$$\dot{\boldsymbol{x}}(t) = \frac{\tau_f - \tau_0}{2}\{\boldsymbol{f}[\boldsymbol{x}(t), \boldsymbol{u}(t)]\} \tag{15}$$

$$\psi_0[\boldsymbol{x}(-1), \tau_0] = \boldsymbol{0} \tag{16}$$

$$\psi_f[\boldsymbol{x}(1), \tau_f] = \boldsymbol{0} \tag{17}$$

$$\boldsymbol{g}[\boldsymbol{u}(t)] \le \boldsymbol{0} \tag{18}$$

Strictly speaking, as a result of the transformation $\tau \to t$ we must adopt new symbols for the variables $\boldsymbol{x}$, $\boldsymbol{u}$, $\boldsymbol{\lambda}$, $\boldsymbol{\mu}$, and $\boldsymbol{\nu}$ and the maps

$$\mathcal{J}(\cdot),\ \mathcal{L}(\cdot),\ \mathcal{M}(\cdot),\ \boldsymbol{f}(\cdot),\ \psi_0(\cdot),\ \psi_f(\cdot)$$

However, for the purpose of brevity, we abuse the notation here and elsewhere and retain these symbols. In this context, Eq. (7) is written as

$$\mathcal{H}(\boldsymbol{x}, \boldsymbol{\lambda}, \boldsymbol{u}) = [(\tau_f - \tau_0)/2]\boldsymbol{\lambda}^T\boldsymbol{f} + [(\tau_f - \tau_0)/2]\mathcal{L} + \boldsymbol{\mu}^T\boldsymbol{g} \tag{19}$$

Let $L_N(t)$ be the Legendre polynomial of degree N on the interval $[-1, 1]$. In the Legendre collocation approximation[11,12] of Eqs. (14–18), we use the LGL points, t_l, $l = 0, \ldots, N$, which are given by

$$t_0 = -1, \qquad t_N = 1$$

and for $1 \le l \le N-1$, t_l are the zeros of $\dot{L}_N$, the derivative of the Legendre polynomial L_N. In the first step of this method, we approximate the continuous variables by Nth degree polynomials of the form

$$\boldsymbol{x}(t) \approx \boldsymbol{x}^N(t) = \sum_{l=0}^{N} \boldsymbol{x}(t_l)\phi_l(t) \tag{20}$$

$$\boldsymbol{u}(t) \approx \boldsymbol{u}^N(t) = \sum_{l=0}^{N} \boldsymbol{u}(t_l)\phi_l(t) \tag{21}$$

where for $l = 0, 1, \ldots, N$

$$\phi_l(t) = \frac{1}{N(N+1)L_N(t_l)} \frac{(t^2-1)\dot{L}_N(t)}{t - t_l}$$

are the Lagrange polynomials of order N. It can be shown that

$$\phi_l(t_k) = \delta_{lk} = \begin{cases} 1 & \text{if} \quad l = k \\ 0 & \text{if} \quad l \ne k \end{cases}$$

From this property of ϕ_l it follows that

$$\boldsymbol{x}^N(t_l) = \boldsymbol{x}(t_l), \; \boldsymbol{u}^N(t_l) = \boldsymbol{u}(t_l) \tag{22}$$

To carry out the second step of the collocation method, we impose the condition that the given approximations satisfy the differential equations exactly at the LGL collocation points. To express the derivative $\dot{\boldsymbol{x}}^N(t)$ in terms of $\boldsymbol{x}^N(t)$ at the collocation points t_k, we differentiate Eq. (20) and evaluate the result at t_k to obtain a matrix multiplication of the following form[11–13]:

$$\dot{\boldsymbol{x}}^N(t_k) = \sum_{l=0}^{N} D_{kl}\boldsymbol{x}(t_l) \tag{23}$$

where D_{kl} are entries of the $(N+1)\times(N+1)$ differentiation matrix $\boldsymbol{D}$

$$\boldsymbol{D} := [D_{kl}] := \begin{cases} \dfrac{L_N(t_k)}{L_N(t_l)} \cdot \dfrac{1}{t_k - t_l} & k \ne l \\ -\dfrac{N(N+1)}{4} & k = l = 0 \\ \dfrac{N(N+1)}{4} & k = l = N \\ 0 & \text{otherwise} \end{cases} \tag{24}$$

To facilitate the NLP formulation, we use the notation

$$\boldsymbol{a}_l := \boldsymbol{x}(t_l), \qquad \boldsymbol{b}_l := \boldsymbol{u}(t_l)$$

to rewrite Eqs. (20) and (21) in the form

$$\boldsymbol{x}^N(t) = \sum_{l=0}^{N} \boldsymbol{a}_l\phi_l(t) \tag{25}$$

$$\boldsymbol{u}^N(t) = \sum_{l=0}^{N} \boldsymbol{b}_l\phi_l(t) \tag{26}$$

For the derivative of the state vector $\boldsymbol{x}(t)$, collocated at the points t_k, we rewrite Eq. (23)

$$\boldsymbol{c}_k = \dot{\boldsymbol{x}}^N(t_k) = \sum_{l=0}^{N} D_{kl}\boldsymbol{a}_l \tag{27}$$

Next, the integral in Eq. (14) is discretized: Substituting Eqs. (25) and (26) in Eq. (14) and using the Gauss–Lobatto integration rule, we obtain

$$\begin{aligned} \mathcal{J}^N(\boldsymbol{a},\boldsymbol{b},\tau_f) &= \mathcal{M}\big[\boldsymbol{x}^N(1),\tau_f\big] + \frac{\tau_f-\tau_0}{2}\int_{-1}^{1}\mathcal{L}[\boldsymbol{x}^N(t),\boldsymbol{u}^N(t)]\,\mathrm{d}t \\ &= \mathcal{M}\big[\boldsymbol{x}^N(1),\tau_f\big] + \frac{\tau_f-\tau_0}{2}\sum_{k=0}^{N}\mathcal{L}\left[\sum_{l=0}^{N}\boldsymbol{a}_l\phi_l(t_k), \sum_{l=0}^{N}\boldsymbol{b}_l\phi_l(t_k)\right]w_k \\ &= \mathcal{M}\big(\boldsymbol{a}_N,\tau_f\big) + \frac{\tau_f-\tau_0}{2}\sum_{k=0}^{N}\mathcal{L}(\boldsymbol{a}_k,\boldsymbol{b}_k)w_k \end{aligned} \tag{28}$$

where the last equality is obtained from $\phi_l(t_k) = \delta_{lk}$. The coefficients are $\boldsymbol{a} = (\boldsymbol{a}_0, \boldsymbol{a}_1, \ldots, \boldsymbol{a}_N)$ and $\boldsymbol{b} = (\boldsymbol{b}_0, \boldsymbol{b}_1, \ldots, \boldsymbol{b}_N)$ and w_k are the weights given by

$$w_k := \{2/[N(N+1)]\}\big\{1/[L_N(t_k)]^2\big\}, \qquad k = 0, 1, \ldots, N$$

The state equations and the initial and terminal state conditions are discretized by first substituting Eqs. (23–26) in Eq. (15) and collocating at the LGL nodes t_k. Using the notation for $\boldsymbol{a}$ and $\boldsymbol{b}$, the state equations are transformed into the following algebraic equations:

$$[(\tau_f - \tau_0)/2]\boldsymbol{f}(\boldsymbol{a}_k,\boldsymbol{b}_k) - \boldsymbol{c}_k = \boldsymbol{0}, \qquad k = 0, \ldots, N$$

where $\boldsymbol{c}_k$ is as defined in Eq. (27) and the initial conditions are

$$\psi_0\big[\boldsymbol{x}^N(-1),\tau_0\big] = \boldsymbol{0} \quad \text{or} \quad \psi_0(\boldsymbol{a}_0,\tau_0) = \boldsymbol{0}$$

The terminal state conditions are

$$\psi_f\big[\boldsymbol{x}^N(1),\tau_f\big] = \boldsymbol{0} \quad \text{or} \quad \psi_f(\boldsymbol{a}_N,\tau_f) = \boldsymbol{0}$$

The control inequality constraints are approximated by

$$\boldsymbol{g}\big[\boldsymbol{u}^N(t_k)\big] \le \boldsymbol{0}, \qquad k = 0, \ldots, N$$

or

$$\boldsymbol{g}(\boldsymbol{b}_k) \le \boldsymbol{0}, \qquad k = 0, \ldots, N$$

To summarize, the optimal control problem in Eqs. (14–18) is approximated by the following nonlinear optimization problem: Find coefficients

$$\boldsymbol{a} = (\boldsymbol{a}_0, \boldsymbol{a}_1, \ldots, \boldsymbol{a}_N), \qquad \boldsymbol{b} = (\boldsymbol{b}_0, \boldsymbol{b}_1, \ldots, \boldsymbol{b}_N)$$

and possibly the final time τ_f to minimize

$$\mathcal{J}^N(\boldsymbol{a},\boldsymbol{b}) = \frac{\tau_f-\tau_0}{2}\sum_{k=0}^{N}\mathcal{L}(\boldsymbol{a}_k,\boldsymbol{b}_k)w_k + \mathcal{M}(\boldsymbol{a}_N,\tau_f) \tag{29}$$

subject to

$$\left(\frac{\tau_f-\tau_0}{2}\right)\boldsymbol{f}(\boldsymbol{a}_k,\boldsymbol{b}_k) - \sum_{l=0}^{N} D_{kl}\boldsymbol{a}_l = \boldsymbol{0}, \qquad k = 0, \ldots, N \tag{30}$$

$$\boldsymbol{g}(\boldsymbol{b}_k) \le \boldsymbol{0}, \qquad k = 0, \ldots, N \tag{31}$$

$$\psi_0(\boldsymbol{a}_0,\tau_0) = \boldsymbol{0} \tag{32}$$

$$\psi_f(\boldsymbol{a}_N,\tau_f) = \boldsymbol{0} \tag{33}$$

From the preceding equations, one can see the simplicity of the method, which retains much of the structure of the continuous problem. This has been achieved by collocating the equations at the LGL points, and except for the differentiation matrix $\boldsymbol{D}$, which relates the different $\boldsymbol{a}_k$, the remaining functions are evaluated only at the LGL points without any dependence on the neighboring points. Note that the state differential constraint Eq. (30), which uses the differentiation matrix to approximate the derivative in the state equations, is very different from the approximation of the state equations achieved by implicit integration rules.

IV. Costate Estimates

In this section, we develop the first-order necessary optimality conditions for the discretized problem described in the preceding section. Without loss of generality we consider the discretization of the problem in the Mayer form. The Lagrangian of the discretized problem Eqs. (29–33) can be written as

$$\tilde{\mathcal{J}}^N = \mathcal{M}\big[\boldsymbol{x}^N(1), \tau_f\big] + \boldsymbol{\nu}_0^T\psi_0\big[\boldsymbol{x}^N(-1), \tau_0\big] + \boldsymbol{\nu}_f^T\psi_f\big[\boldsymbol{x}^N(1), \tau_f\big] + \sum_{i=0}^{N}\left(\tilde{\boldsymbol{\lambda}}_i^T\left\{\frac{(\tau_f-\tau_0)}{2}\boldsymbol{f}[\boldsymbol{x}(t_i),\boldsymbol{u}(t_i)] - \dot{\boldsymbol{x}}(t_i)\right\} + \tilde{\boldsymbol{\mu}}_i\boldsymbol{g}[\boldsymbol{u}(t_i)]\right) \tag{34}$$

where $\tilde{\boldsymbol{\lambda}}_i$ and $\tilde{\boldsymbol{\mu}}_i$ are the Lagrange multipliers associated with the equality (i.e., discretized state equations) and inequality (i.e., discretized control) constraints, respectively.

In terms of the notation introduced in the preceding section, we have

$$\tilde{\mathcal{J}}^N = \mathcal{M}(\boldsymbol{a}_N, \tau_f) + \boldsymbol{\nu}_0^T\psi_0(\boldsymbol{a}_0, \tau_0) + \boldsymbol{\nu}_f^T\psi_f(\boldsymbol{a}_N, \tau_f) + \sum_{i=0}^{N}\left\{\tilde{\boldsymbol{\lambda}}_i^T\left[\frac{(\tau_f-\tau_0)}{2}\boldsymbol{f}_i - \boldsymbol{c}_i\right] + \tilde{\boldsymbol{\mu}}_i^T\boldsymbol{g}_i\right\} \tag{35}$$

where

$$\boldsymbol{f}_i = \boldsymbol{f}[\boldsymbol{x}(t_i), \boldsymbol{u}(t_i)] = \boldsymbol{f}(\boldsymbol{a}_i, \boldsymbol{b}_i), \qquad \boldsymbol{g}_i = \boldsymbol{g}[\boldsymbol{u}(t_i)] = \boldsymbol{g}(\boldsymbol{b}_i)$$

The KKT first-order necessary conditions are[21]

$$\frac{\partial\tilde{\mathcal{J}}^N}{\partial\boldsymbol{a}_k} = \boldsymbol{0}, \qquad \frac{\partial\tilde{\mathcal{J}}^N}{\partial\boldsymbol{b}_k} = \boldsymbol{0}, \qquad \frac{\partial\tilde{\mathcal{J}}^N}{\partial\tau_f} = 0 \qquad \text{for} \quad k = 0,\ldots,N \tag{36}$$

$$\tilde{\boldsymbol{\mu}}_k \geq \boldsymbol{0}, \qquad \tilde{\boldsymbol{\mu}}_k^T\boldsymbol{g}_k = 0 \tag{37}$$

First consider the interior state variables $(\boldsymbol{a}_1,\ldots,\boldsymbol{a}_{N-1})$:

$$\frac{\partial\tilde{\mathcal{J}}^N}{\partial\boldsymbol{a}_k} = \frac{\partial}{\partial\boldsymbol{a}_k}\left\{\sum_{i=0}^{N}\tilde{\boldsymbol{\lambda}}_i^T\left[\frac{(\tau_f-\tau_0)}{2}\boldsymbol{f}_i - \boldsymbol{c}_i\right]\right\} \tag{38}$$

Because the function $\boldsymbol{f}$ is collocated only at the points t_i, the term $\boldsymbol{f}_i$ does not depend on the adjacent points t_i; hence,

$$\frac{\partial\boldsymbol{f}_i}{\partial\boldsymbol{a}_k} = \delta_{ik}\frac{\partial\boldsymbol{f}_i}{\partial\boldsymbol{a}_k} \tag{39}$$

Therefore,

$$\frac{\partial}{\partial\boldsymbol{a}_k}\left\{\sum_{i=0}^{N}\tilde{\boldsymbol{\lambda}}_i^T\left[\frac{(\tau_f-\tau_0)}{2}\boldsymbol{f}_i\right]\right\} = \frac{(\tau_f-\tau_0)}{2}\left(\frac{\partial\boldsymbol{f}_k}{\partial\boldsymbol{a}_k}\right)^T\tilde{\boldsymbol{\lambda}}_k$$

For the partials of the state derivatives, $\partial\dot{\boldsymbol{x}}_i/\partial\boldsymbol{a}_k = \partial\boldsymbol{c}_i/\partial\boldsymbol{a}_k$, a more complicated expression is obtained because the differentiation matrix $\boldsymbol{D}$ relates the different components of $\boldsymbol{a}_k$:

$$\frac{\partial\boldsymbol{c}_i}{\partial\boldsymbol{a}_k} = \frac{\partial}{\partial\boldsymbol{a}_k}\sum_{l=0}^{N}D_{il}\boldsymbol{a}_l = D_{ik}\boldsymbol{I}^n \tag{40}$$

where $\boldsymbol{I}^n$ is the n-dimensional identity matrix and n is the dimension of the state vector. Note that D_{ik} is a scalar multiplied to $\boldsymbol{I}^n$. Thus, part of expression (38) simplifies to

$$\frac{\partial}{\partial\boldsymbol{a}_k}\left[\sum_{i=0}^{N}\tilde{\boldsymbol{\lambda}}_i^T\boldsymbol{c}_i\right] = \sum_{i=0}^{N}D_{ik}\tilde{\boldsymbol{\lambda}}_i \tag{41}$$

From the following expressions for w_i and D_{ik}:

$$w_k := \frac{2}{N(N+1)}\frac{1}{[L_N(t_k)]^2}, \qquad k = 0,1,\ldots,N$$

$$D_{ik} = \frac{L_N(t_i)}{L_N(t_k)}\frac{1}{t_i - t_k}, \qquad i \neq k$$

we have

$$\begin{aligned} w_iD_{ik} &= \frac{2}{N(N+1)}\frac{1}{[L_N(t_i)]^2}\frac{L_N(t_i)}{L_N(t_k)}\cdot\frac{1}{t_i-t_k} \\ &= \frac{2}{N(N+1)}\frac{1}{[L_N(t_i)]L_N(t_k)}\frac{1}{t_i-t_k} \\ &= \left\{\frac{2}{N(N+1)}\frac{1}{[L_N(t_k)]^2}\right\}\frac{L_N(t_l)}{L_N(t_i)}\frac{-1}{t_k-t_i} \\ &= -w_kD_{ki} \end{aligned} \tag{42}$$

From the expression $D_{ik} = -(w_k/w_i)D_{ki}$, (note that $D_{ii} = 0$ for $i \neq 0, N$) and Eqs. (39) and (41), the following is obtained for $k = 1,\ldots,N-1$:

$$\frac{\partial\tilde{\mathcal{J}}^N}{\partial\boldsymbol{a}_k} = \frac{(\tau_f-\tau_0)}{2}\left(\frac{\partial\boldsymbol{f}_k}{\partial\boldsymbol{a}_k}\right)^T\tilde{\boldsymbol{\lambda}}_k + w_k\sum_{i=0}^{N}D_{ki}\left(\frac{\tilde{\boldsymbol{\lambda}}_i}{w_i}\right) = \boldsymbol{0} \tag{43}$$

Dividing Eq. (43) by w_k and comparing it to the continuous costate equations collocated at the LGL point, as

$$\dot{\boldsymbol{\lambda}}(t_k) = \sum_{i=0}^{N}D_{ki}\boldsymbol{\lambda}(t_i) = -\left(\frac{\tau_f-\tau_0}{2}\right)\left(\frac{\partial\boldsymbol{f}}{\partial\boldsymbol{x}_k}\right)^T(t_k)\boldsymbol{\lambda}(t_k) \qquad \text{for} \quad k = 1,\ldots,N-1 \tag{44}$$

we have the following result:

$$\boldsymbol{\lambda}(t_k) = \tilde{\boldsymbol{\lambda}}_k/w_k, \qquad k = 1,\ldots,N-1 \tag{45}$$

In other words, the KKT multipliers divided by the LGL weights are the same as the costates at the interior nodes.

To find an expression for $\partial/\partial\boldsymbol{a}_N$, we proceed as before and take the partial of every term in Eq. (35) with respect to $\boldsymbol{a}_N$,

$$\frac{\partial\tilde{\mathcal{J}}^N}{\partial\boldsymbol{a}_N} = \frac{\partial\mathcal{M}}{\partial\boldsymbol{a}_N} + \left(\frac{\partial\psi_f}{\partial\boldsymbol{a}_N}\right)^T\boldsymbol{\nu}_f + \frac{(\tau_f-\tau_0)}{2}\left(\frac{\partial\boldsymbol{f}_N}{\partial\boldsymbol{a}_N}\right)^T\tilde{\boldsymbol{\lambda}}_N - \sum_{i=0}^{N}D_{iN}\tilde{\boldsymbol{\lambda}}_i = \boldsymbol{0} \tag{46}$$

Using the relationship

$$D_{iN} = -(w_N/w_i)D_{Ni}, \qquad i \neq N$$

and adding $2D_{NN}\tilde{\boldsymbol{\lambda}}_N$ to both sides of Eq. (46), we obtain

$$\left[\frac{(\tau_f-\tau_0)}{2}\left(\frac{\partial\boldsymbol{f}_N}{\partial\boldsymbol{a}_N}\right)^T\tilde{\boldsymbol{\lambda}}_N + w_N\sum_{i=0}^{N}D_{Ni}\frac{\tilde{\boldsymbol{\lambda}}_i}{w_i}\right] = 2D_{NN}\tilde{\boldsymbol{\lambda}}_N - \frac{\partial\mathcal{M}}{\partial\boldsymbol{a}_N} - \left(\frac{\partial\psi_f}{\partial\boldsymbol{a}_N}\right)^T\boldsymbol{\nu}_f \tag{47}$$

From the definition of w_N and D_{NN}

$$w_N = \frac{2}{N(N+1)}\frac{1}{L_N(t_N)^2}, \qquad L_N(t_N = 1) = 1$$

$$D_{NN} = [N(N+1)]/4$$

we have

$$2D_{NN} = 1/w_N$$

Therefore, Eq. (47) can be written as

$$w_N\left[\frac{(\tau_f - \tau_0)}{2}\left(\frac{\partial \boldsymbol{f}_N}{\partial \boldsymbol{a}_N}\right)^T \frac{\tilde{\boldsymbol{\lambda}}_N}{w_N} + \sum_{i=0}^{N} D_{Ni}\frac{\tilde{\boldsymbol{\lambda}}_i}{w_i}\right]$$

$$= \frac{\tilde{\boldsymbol{\lambda}}_N}{w_N} - \frac{\partial \mathcal{M}}{\partial \boldsymbol{a}_N} - \left(\frac{\partial \boldsymbol{\psi}_f}{\partial \boldsymbol{a}_N}\right)^T \boldsymbol{\nu}_f \tag{48}$$

This equation is the combination of the costate equation collocated at t_N and the final time transversality condition for $\boldsymbol{\lambda}(t_N)$. If we choose $\boldsymbol{\nu}_f$ to satisfy the final time transversality condition

$$\frac{\tilde{\boldsymbol{\lambda}}_N}{w_N} = \frac{\partial \mathcal{M}}{\partial \boldsymbol{a}_N} + \left(\frac{\partial \boldsymbol{\psi}_f}{\partial \boldsymbol{a}_N}\right)^T \boldsymbol{\nu}_f \tag{49}$$

we also get

$$\frac{(\tau_f - \tau_0)}{2}\left(\frac{\partial \boldsymbol{f}_N}{\partial \boldsymbol{a}_N}\right)^T \frac{\tilde{\boldsymbol{\lambda}}_N}{w_N} + \sum_{i=0}^{N} D_{Ni}\frac{\tilde{\boldsymbol{\lambda}}_i}{w_i} = \boldsymbol{0} \tag{50}$$

which is exactly the costate differential equation collocated at t_N. Therefore, we have

$$\boldsymbol{\lambda}(t_N) = \tilde{\boldsymbol{\lambda}}_N / w_N$$

The case of initial time condition can be shown in a similar fashion:

$$\frac{(\tau_f - \tau_0)}{2}\left(\frac{\partial \boldsymbol{f}_0}{\partial \boldsymbol{a}_0}\right)^T \frac{\tilde{\boldsymbol{\lambda}}_0}{w_0} + \sum_{i=0}^{N} D_{0i}\frac{\tilde{\boldsymbol{\lambda}}_i}{w_i} = \boldsymbol{0} \tag{51}$$

$$\frac{\tilde{\boldsymbol{\lambda}}_0}{w_0} = -\left(\frac{\partial \boldsymbol{\psi}_0}{\partial \boldsymbol{a}_0}\right)^T \boldsymbol{\nu}_0 \tag{52}$$

For the control optimality conditions, taking the partial of $\tilde{\mathcal{J}}^N$ with respect to $\boldsymbol{b}_i$ yields

$$\frac{\partial \tilde{\mathcal{J}}^N}{\partial \boldsymbol{b}_i} = \left[\frac{(\tau_f - \tau_0)}{2}\left(\frac{\partial \boldsymbol{f}_i}{\partial \boldsymbol{b}_i}\right)^T \tilde{\boldsymbol{\lambda}}_i + \left(\frac{\partial \boldsymbol{g}_i}{\partial \boldsymbol{b}_i}\right)^T \tilde{\boldsymbol{\mu}}_i\right] = \boldsymbol{0}$$

$$i = 0, \ldots, N \tag{53}$$

and $\tilde{\boldsymbol{\mu}}_i^T \boldsymbol{g}_i = 0$ and $\tilde{\boldsymbol{\mu}}_i \geq \boldsymbol{0}$. Dividing Eq. (53) by w_i, it is clear that it yields the discrete analog of Eq. (8) with the discrete Hamiltonian defined as [see Eq. (19)]

$$\mathcal{H}[\boldsymbol{x}(t_i), \boldsymbol{\lambda}(t_i), \boldsymbol{u}(t_i)] = [(\tau_f - \tau_0)/2]\boldsymbol{\lambda}^T(t_i)\boldsymbol{f}_i + \boldsymbol{\mu}(t_i)^T \boldsymbol{g}_i \tag{54}$$

Therefore,

$$\boldsymbol{\mu}(t_i) = \tilde{\boldsymbol{\mu}}_i / w_i \tag{55}$$

To summarize, the continuous Lagrange multipliers for the optimal control problem can be obtained at the collocation points by simply dividing the KKT multipliers by the LGL weights. The optimality of the direct solution can now be checked in several ways; for example, the optimal control obtained from the costates can be compared with the control solution from the direct method.

V. Numerical Examples

To illustrate the theory presented in the preceding sections, we select two examples: one from a recent paper by Herman and Conway,[7] and another from the text by Bryson and Ho.[22]

Example 1

Consider the problem of determining the optimal trajectory and the thrust steering vector to transfer a rocket from an initial orbit to a target orbit in fixed time.[7] The state variables are the radial distance r, the true anomaly θ, the radial component of velocity u, and the tangential component of velocity v. The control variable is the thrust steering angle measured from the local horizontal ϵ. The optimal control problem is formulated as finding $\epsilon(\tau)$ to maximize the final energy. Therefore, the cost function is defined as

$$\mathcal{J} = \mathcal{M}(\tau_f) = -\left\{\tfrac{1}{2}\left[u(\tau_f)^2 + v(\tau_f)^2\right] - [1/r(\tau_f)]\right\} \tag{56}$$

subject to the canonical equations of motion

$$\frac{dr}{d\tau} = u \tag{57}$$

$$\frac{d\theta}{d\tau} = \frac{v}{r} \tag{58}$$

$$\frac{du}{d\tau} = \frac{v^2}{r} - \frac{1}{r^2} + A(\tau)\sin\epsilon \tag{59}$$

$$\frac{dv}{d\tau} = -\frac{uv}{r} + A(\tau)\cos\epsilon \tag{60}$$

where $A(\tau) = 0.01$. The initial conditions for this problem are

$$r(0) = 1.1, \qquad \theta(0) = 0 \tag{61}$$

$$u(0) = 0, \qquad v(0) = 1\big/\sqrt{1.1} \tag{62}$$

The final time is fixed at $\tau_f = 50$, and the final states are free.

In the first set of simulations, the Legendre pseudospectral method as summarized in Eqs. (29–33) was utilized to formulate the NLP for minimizing the cost function (56) subject to the approximations of the state equations (57–60) and the initial conditions (61–62).

The NLP problem for the state and control variables was solved for 64 LGL points. All of the computations were performed using MATLAB® 5.2 and NPSOL as the NLP solver.[23] We did not choose *constr.m* or *fmincon.m*, the constrained minimization programs in the MATLAB Optimization Toolbox, because these solvers converged to inaccurate solutions and did not provide the correct multipliers. For the initial guess, a solution from the numerical integration of the state equations with a constant steering of 0.001 rad was used. The optimal trajectory is shown in Fig. 1, and the optimal velocity components u and v are shown in Fig. 2. Our optimal cost function of -0.09512 is in perfect agreement with -0.09512 reported in Ref. 7. We should note that for 64 LGL points the size of the NLP variable for our problem is $5 \times 64 = 320$. In Ref. 7, where

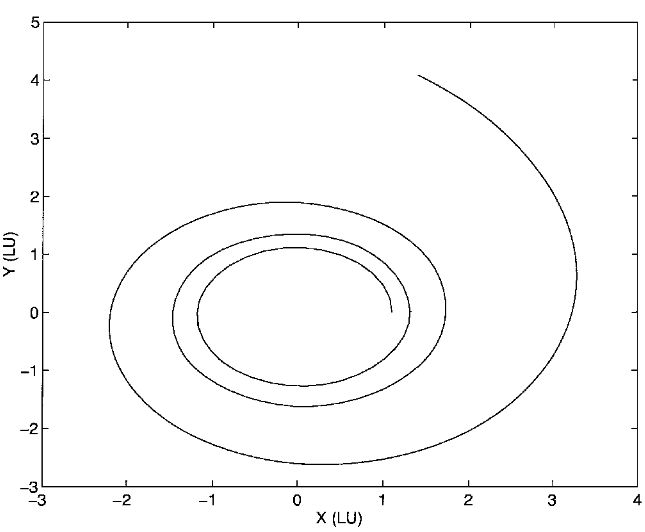

Fig. 1 Optimizing trajectory.

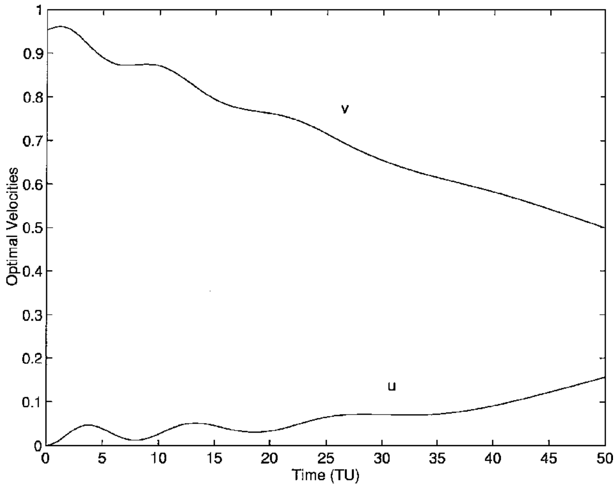

Fig. 2 Optimal states u and v vs time.

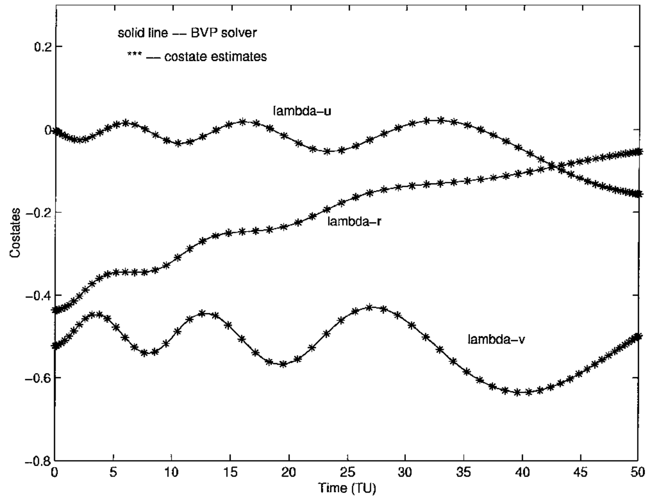

Fig. 3 Costates λ_r, λ_u, and λ_v vs time.

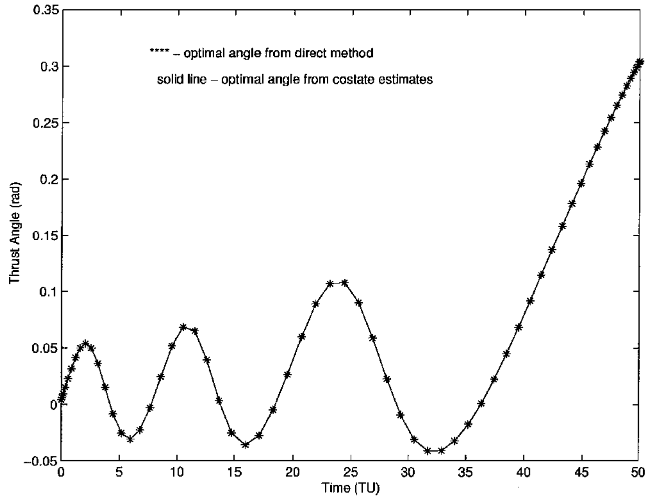

Fig. 4 Time history of the optimal thrust steering angle.

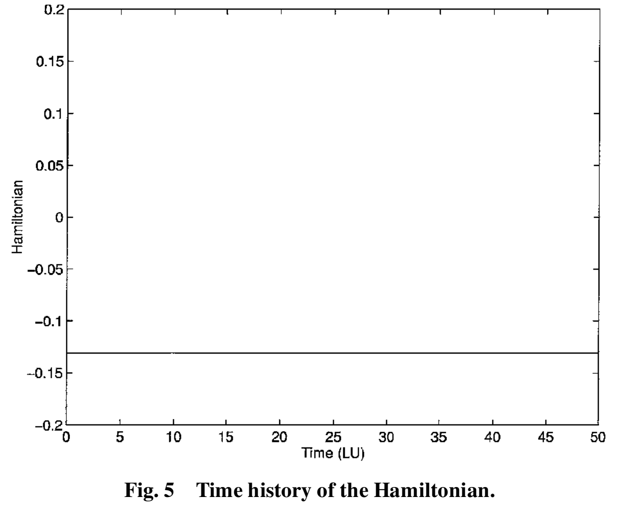

Fig. 5 Time history of the Hamiltonian.

the same problem is solved for 30 subintervals using a fifth-degree Gauss–Lobatto collocation method, the size of the NLP variable is 365. Therefore, we obtain comparably accurate results to the ones in Ref. 7 with fewer number of collocation points.

The optimality of our solution may now be checked in several ways. The terminal transversality conditions are given by

$$\lambda_r(\tau_f) = -1\big/[r(\tau_f)]^2 \tag{63}$$

$$\lambda_\theta(\tau_f) = 0 \tag{64}$$

$$\lambda_u(\tau_f) = -u(\tau_f) \tag{65}$$

$$\lambda_v(\tau_f) = -v(\tau_f) \tag{66}$$

$$\mathcal{H}(\tau_f) = -\nu_f \tag{67}$$

where ν_f is the multiplier associated with the final time constraint, $\tau_f - 50 = 0$. Thus, the costates at τ_f should be

$$\boldsymbol{\lambda}(\tau_f) = [-0.0537, 0, -0.1566, -0.4986]^T \tag{68}$$

The costates estimates obtained by using the values of the KKT multipliers from NPSOL, the NLP solver in our problem, are computed at final time $\tau_f = 50$ to be

$$\boldsymbol{\lambda}(\tau_f) = [-0.0536, 2.776 \times 10^{-4}, -0.1565, -0.4984]^T \tag{69}$$

Clearly, this is in excellent agreement with Eq. (68).

The costates' history is shown in Fig. 3. The results from the costate estimates are contrasted against the results from a boundary value problem (BVP) solver, and the agreement between the two sets of data is excellent. Additional checks on the solution may be proposed as follows. The optimal control must satisfy

$$\tan\epsilon = \lambda_u \cos(\epsilon) - \lambda_v \sin(\epsilon) = 0 \tag{70}$$

In Fig. 4, the control from the direct method is compared with the one obtained from the costate estimates and Eq. (70). It is obvious that the results are in excellent agreement. The Hamiltonian for this autonomous problem should be a constant, and our computations of the Hamiltonian shown in Fig. 5 independently demonstrate the validity of our results.

Example 2

In this example, the dynamics are nearly the same as that of Example 1, with one exception

$$A(\tau) = T/(m_0 - |\dot{m}|\tau) \tag{71}$$

where m_0 is the initial mass and $\dot{m}$ is the constant fuel consumption rate. The optimal control problem, as discussed in Refs. 22 and 24, is to find the control $\epsilon(\tau)$ that maximizes the final radius at $\tau_f = 3.32$. Thus,

$$\mathcal{J} = -r(t_N)$$

and the state dynamics are the same as that of the preceding example. The boundary conditions are

$$r(0) = 1.0, \qquad \theta(0) = 0 \tag{72}$$

$$u(0) = 0, \qquad v(0) = 1.0 \tag{73}$$

$$u(\tau_f) = 0, \qquad v(\tau_f) - \sqrt{1/r(\tau_f)} = 0 \tag{74}$$

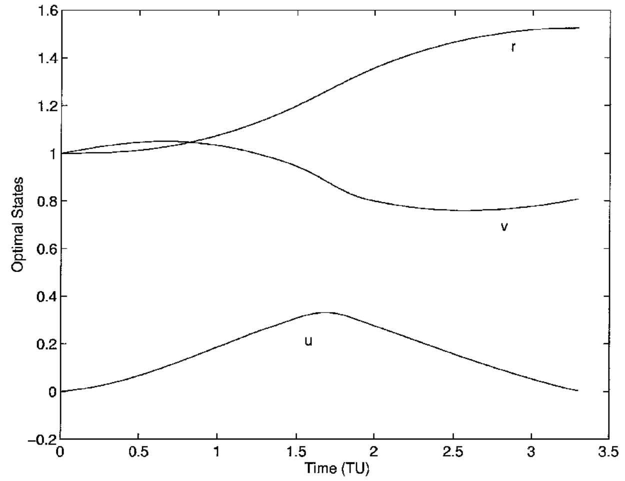

Fig. 6 Optimal states *r*, *u*, and *v* vs time.

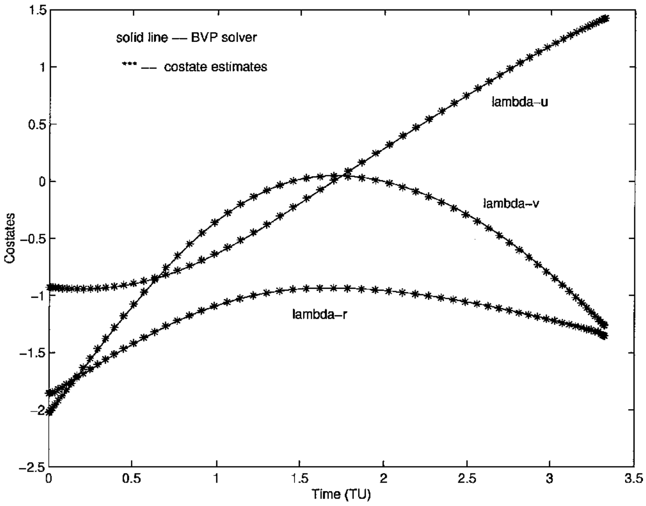

Fig. 7 Costates λ_r, λ_u, and λ_v.

The normalized constants in this problem are $m_0 = 1.0$, $T = 0.1405$, $\tau_f = 3.32$, and $|\dot{m}| = 0.0749$.

For the simulations, the number of LGL points was 64 as in Example 1. For the initial guess, a solution obtained from numerically integrating the differential equations with a constant value for the control $\epsilon(\tau) = 0.001$ rad was used.

Figure 6 shows a plot of the converged optimal states. These plots are not provided in Ref. 22, but the performance index we obtained was $r(t_N) = 1.525$, which is in good agreement with Ref. 24. The results for the costates are shown in Fig. 7. As in the preceding example, the costates estimates are shown against the results from an indirect method (a BVP solver), and the accuracy of the estimates is confirmed from their excellent match with the indirect results. In Ref. 22, the optimal control history $\epsilon(\tau)$ is displayed as a direction field, but Moyer and Pinkham in Ref. 24 provide a profile for $\epsilon(\tau)$. Our results from the direct method, shown in Fig. 8, agree well with their results. In Fig. 8, we also shows the controls obtained from Eq. (70) from the BVP solver. Note that the sharp variation in ϵ around $\tau = 1.5$ is captured by the method quite adequately.

Finally, note that, in general, there is often a difference between a method and its implementation.[1] For example, the LGL points are theoretically the roots of the derivatives of the Legendre polynomials. Computationally, we determine these points not by differentiating the polynomials and solving for the roots, but by the use of advanced numerical linear algebra techniques that yield faster and more accurate results.[25] In similar vein, the costate estimates obtained simply by using Eq. (45) tend to be noisy with a characteristic spectrum that needs to be filtered. The details of such implementation techniques are beyond the scope of the present paper, but they are discussed elsewhere.[17]

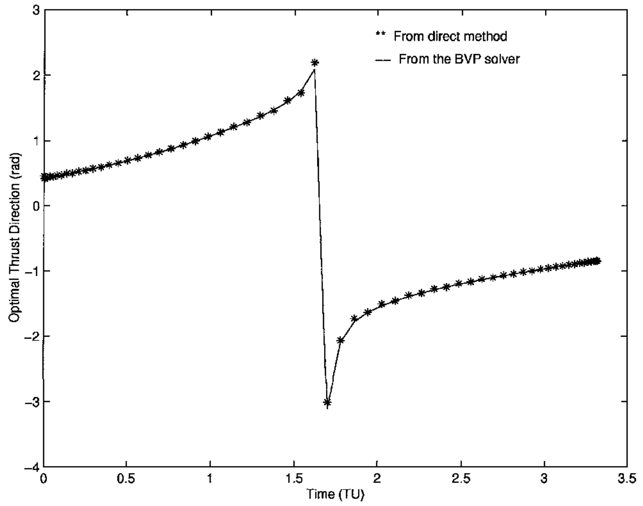

Fig. 8 Graphs of ϵ from the direct method vs ϵ from the optimality condition.

VI. Conclusions

A Legendre spectral collocation method has been presented for costate estimates based on the results from the direct optimization. It is proved that the costates at the LGL points are equal to the Lagrange multipliers of the approximate NLP problem divided by the LGL weights. As demonstrated by the numerical examples, the costates estimated by our method are in excellent agreement with those obtained by an indirect method. Thus, the optimality of the direct solution may be readily checked in several different ways. The method is easy to implement as it preserves the structure of the original optimal control problem.

Acknowledgments

Funding for this research was provided in part by the Naval Postgraduate School and the Naval Space Command. We thank Craig Baldwin for supporting this research and William Gragg for providing the numerical codes for computing the Legendre–Gauss–Lobatto points. We are indebted to Philip Gill and Michael Gertz for not only helping us with NPSOL but also for their patience in answering our various questions. Finally, we express our deepest gratitude to T. Thorvaldsen of Charles Stark Draper Laboratory for his elegant design of a digital filter.

References

[1] Betts, J. T., "Survey of Numerical Methods for Trajectory Optimization," *Journal of Guidance, Control, and Dynamics*, Vol. 21, No. 2, 1998, pp. 193–207.

[2] Hull, D. G., "Conversion of Optimal Control Problems into Parameter Optimization Problems," *Journal of Guidance, Control, and Dynamics*, Vol. 20, No. 1, 1997, pp. 57–60.

[3] Seywald, H., "Trajectory Optimization Based on Differential Inclusion," *Journal of Guidance, Control, and Dynamics*, Vol. 17, No. 3, 1994, pp. 480–487.

[4] Fahroo, F., and Ross, I. M., "Computational Optimal Control by Spectral Collocation with Differential Inclusion," *Proceedings of the 1999 Flight Mechanics Symposium*, NASA CP-1999-209235, 1999, pp. 185–200.

[5] 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.

[6] von Stryk, O., "Numerical Solution of Optimal Control Problems by Direct Collocation," *Optimal Control*, edited by R. Bulirsch, A. Miele and J. Stoer, Birkhäuser Verlag, Basel, 1993, pp. 129–143.

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

[8] Vlassenbroeck, J., and Van Doreen, R., "A Chebyshev Technique for Solving Nonlinear Optimal Control Problems," *IEEE Transactions on Automatic Control*, Vol. 33, No. 4, 1988, pp. 333–340.

[9] Vlassenbroeck, J., "A Chebyshev Polynomial Method for Optimal Control with State Constraints," *Automatica*, Vol. 24, 1988, pp. 499–506.

[10] Fahroo, F., and Ross, I. M., "Direct Trajectory Optimization by a Chebyshev Pseudospectral Method," *Proceedings of the American Control Conference*, 2000, pp. 3860–3864.

[11]Gottlieb, D., Hussaini, M. Y., and Orszag, S. A., "Theory and Applications of Spectral Methods," *Spectral Methods for PDEs*, Society for Industrial and Applied Mathematics, Philadelphia, 1984.

[12]Canuto, C., Hussaini, M. Y., Quarteroni, A., and Zang, T. A., *Spectral Methods in Fluid Dynamics*, Springer-Verlag, New York, 1988.

[13]Elnagar, J., Kazemi, M. A., and Razzaghi, M., "The Pseudospectral Legendre Method for Discretizing Optimal Control Problems," *IEEE Transactions on Automatic Control*, Vol. 40, No. 10, 1995, pp. 1793–1796.

[14]Elnagar, J., and Razzaghi, M., "A Collocation-Type Method for Linear Quadratic Optimal Control Problems," *Optimal Control Applications and Methods*, Vol. 18, 1997, pp. 227–235.

[15]Fahroo, F., and Ross, I. M., "A Spectral Patching Method for Direct Trajectory Optimization," *Journal of Astronautical Sciences* (to be published).

[16]Davis, P. J., and Rabinowitz, P., *Methods of Numerical Integration*, Academic, New York, 1977.

[17]Fahroo, F., and Ross, I. M., "Trajectory Optimization by Indirect Spectral Collocation Methods," AIAA Paper 2000-4028, Aug. 2000.

[18]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. 3, 1992, pp. 994–1002.

[19]Martell, C. A., and Lawton, J.A., "Adjoint Variable Solutions via an Auxiliary Optimization Problem," *Journal of Guidance, Control, and Dynamics*, Vol. 18, No. 6, 1995, pp. 1267–1272.

[20]Seywald, H., and Kumar, R. R., "Method for Automatic Costate Calculation," *Journal of Guidance, Control, and Dynamics*, Vol. 19, No. 6, 1996, pp. 1252–1261.

[21]Fletcher, R., *Practical Methods of Optimization*, 2nd ed., Wiley, Essex, England, U.K., 1987.

[22]Bryson, A. E., and Ho, Y. C., *Applied Optimal Control*, Hemisphere, New York, 1975.

[23]Gill, P. E., Murray, W., Saunders, M. A., and Wright, M. A., "User's Guide to NPSOL 5.0: A Fortran Package for Nonlinear Programming," TR SOL 86-1, Stanford Optimization Lab., Stanford Univ., July 1998.

[24]Moyer, H. G., and Pinkham, G., "Several Trajectory Optimization Techniques, Part II: Applications," *Computing Methods in Optimization Problems*, edited by A. V. Balakrishnan and L. W. Neustadt, Academic, New York, 1964.

[25]Golub, G. H., "Some Modified Matrix Eigenvalue Problems," *SIAM Review*, Vol. 15, No. 2, 1973, pp. 318–334.