Lecture Notes in Control and Information Sciences, Vol. 295, Springer-Verlag, 2003

Legendre Pseudospectral Approximations of Optimal Control Problems

I. Michael Ross[1] and Fariba Fahroo[2]

[1] Department of Aeronautics and Astronautics, Code AA/Ro, Naval Postgraduate School, Monterey, CA 93943 imross@nps.navy.mil
[2] Department of Applied Mathematics, Code MA/Ff, Naval Postgraduate School, Monterey, CA 93943 ffahroo@nps.navy.mil

Summary. We consider nonlinear optimal control problems with mixed state-control constraints. A discretization of the Bolza problem by a Legendre pseudospectral method is considered. It is shown that the operations of discretization and dualization are not commutative. A set of Closure Conditions are introduced to commute these operations. An immediate consequence of this is a Covector Mapping Theorem (CMT) that provides an order-preserving transformation of the Lagrange multipliers associated with the discretized problem to the discrete covectors associated with the optimal control problem. A natural consequence of the CMT is that for pure state-constrained problems, the dual variables can be easily related to the D-form of the Lagrangian of the Hamiltonian. We demonstrate the practical advantage of our results by numerically solving a state-constrained optimal control problem without deriving the necessary conditions. The costates obtained by an application of our CMT show excellent agreement with the exact analytical solution.

1 Introduction

Many problems in control theory can be formulated as optimal control problems [5]. From a control engineer's perspective, it is highly desirable to obtain feedback solutions to complex nonlinear optimal control problems. Although the Hamilton-Jacobi-Bellman (HJB) equations provide a framework for this task, they suffer from well-known fundamental problems [1, 3, 5], such as the nonsmoothness of the value function and the "curse of dimensionality". The alternative framework of the Minimum Principle, while more tractable from a control-theoretic point of view, generates open-loop controls if it can be solved at all. The Minimum-Principle approach is also beset with fundamental numerical problems due to the fact that the costates are adjoint to the state perturbation equations [3]. In other words, the Hamiltonian generates a numerically sensitive boundary value problem that may produce such wild trajectories as to exceed the numerical range of the computer [3]. To overcome this difficulty, direct methods have been employed to solve complex optimal control problems arising in engineering applications [2]. While the theoretical properties of Eulerian methods are widely studied [5, 12], they are not practical due to their linear ($O(h)$) convergence rate. On the other hand, collocation methods are practical and widely used [2], but not much can be said about the optimality of the result since these methods do not tie the resulting solutions to either the Minimum Principle or HJB theory. In fact, the popular Hermite-Simpson collocation method and even some Runge-Kutta methods do not converge to the solution of the optimal control problem [10]. This is because an N^{th}-order integration scheme for the differential equations does not necessarily lead to an N^{th}-order approximation scheme for the dual variables. That is, discretization and dualization do not necessarily commute [14]. By imposing additional conditions on the coefficients of Runge-Kutta schemes, Hager[10] was able to transform the adjoint system of the discretized problem to prove the preservation of the order of approximation. Despite this breakthrough, the controls in such methods converge more slowly than the states or the adjoints. This is because, the controls are implicitly approximated to a lower order of accuracy (typically piecewise linear functions) in the discrete time interval.

In this paper, we consider the pseudospectral (PS) discretization of constrained nonlinear optimal control problems with a Bolza cost functional[6, 8, 9]. PS methods differ from many of the traditional discretization methods in the sense that the focus of the approximation is on the tangent bundle than on the differential equation[15]. In this sense, they most closely resemble finite element methods but offer a far more impressive convergence rate known as spectral accuracy[17]. For example, for smooth problems, spectral accuracy implies an exponential convergence rate. We show that the discretization of the constrained Bolza problem by an N^{th}-order Legendre PS method does not lead to an N^{th}-order approximation scheme for the dual variables as previously presumed[7, 9]. However, unlike Hager's Runge-Kutta methods, no conditions on the coefficients of the Legendre polynomials can be imposed to overcome this barrier. Fortunately, a set of simple "closure conditions," that we introduce in this paper, can be imposed on the discrete primal-dual variables so that a linear diagonal transformation of the constrained Lagrange multipliers of the discrete problem provides a consistent approximation to the discrete covectors of the Bolza problem. This is the Covector Mapping Theorem (CMT). For pure state-constrained control problems, the CMT naturally provides a discrete approximation to the costates associated with the so-called D-form of the Lagrangian of the Hamiltonian[11]. This implies that the order of the state-constraint is not a limiting factor and that the interior point constraint at the junction of the state constraint is not explicitly imposed. More importantly, the jump conditions are automatically approximated as a consequence of the CMT. These sets of results offer an enormously practical advantage over other methods and are demonstrated by a numerical example.

2 Problem Formulation

We consider the following formulation of an autonomous, mixed state-control constrained Bolza optimal control problem with possibly free initial and terminal times:

Problem B

Determine the state-control function pair, $[\tau_0, \tau_f] \ni \tau \mapsto \{\boldsymbol{x} \in \mathbb{R}^{N_x}, \boldsymbol{u} \in \mathbb{R}^{N_u}\}$ and possibly the "clock times," τ_0 and τ_f, that minimize the Bolza cost functional,

$$J[\boldsymbol{x}(\cdot), \boldsymbol{u}(\cdot), \tau_0, \tau_f] = E(\boldsymbol{x}(\tau_0), \boldsymbol{x}(\tau_f), \tau_0, \tau_f) + \int_{\tau_0}^{\tau_f} F(\boldsymbol{x}(\tau), \boldsymbol{u}(\tau))\, d\tau \tag{1}$$

subject to the state dynamics,

$$\dot{\boldsymbol{x}}(\tau) = \mathbf{f}(\boldsymbol{x}(\tau), \boldsymbol{u}(\tau)) \tag{2}$$

end-point conditions,

$$\mathbf{e}\big(\boldsymbol{x}(\tau_0), \boldsymbol{x}(\tau_f), \tau_0, \tau_f\big) = \mathbf{0} \tag{3}$$

and mixed state-control path constraints,

$$\mathbf{h}(\boldsymbol{x}(\tau), \boldsymbol{u}(\tau)) \leq \mathbf{0} \tag{4}$$

Assumptions and Notation

For the purpose of brevity, we will make some assumptions that are often not necessary in a more abstract setting. It is assumed the functions $E : \mathbb{R}^{N_x} \times \mathbb{R}^{N_x} \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$, $F : \mathbb{R}^{N_x} \times \mathbb{R}^{N_u} \to \mathbb{R}$, $\mathbf{f} : \mathbb{R}^{N_x} \times \mathbb{R}^{N_u} \to \mathbb{R}^{N_x}$, $\mathbf{e} : \mathbb{R}^{N_x} \times \mathbb{R}^{N_x} \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}^{N_e}$, $\mathbf{h} : \mathbb{R}^{N_x} \times \mathbb{R}^{N_u} \to \mathbb{R}^{N_h}$ are continuously differentiable with respect to their arguments. It is assumed that a feasible solution, and hence an optimal solution exists in an appropriate Sobolev space, the details of which are ignored. In order to apply the first-order optimality conditions, additional assumptions on the constraint set are necessary. Throughout the rest of the paper, such constraint qualifications are implicitly assumed. The Lagrange multipliers discussed in the rest of this paper are all assumed to be nontrivial and regular. The symbol $N_{(\cdot)}$ with a defining subscript is an element of the Natural numbers $\mathbb{N}$. Nonnegative orthants are denoted by $\mathbb{R}_+^{N_h}$. The shorthand $\mathbf{h}[\tau]$ denotes $\mathbf{h}(\boldsymbol{x}(\tau), \boldsymbol{u}(\tau))$. By a somewhat minor abuse of notation, we let $\mathbf{h}_k$ denote $\mathbf{h}^N[\tau_k] = \mathbf{h}(\boldsymbol{x}^N(\tau_k), \boldsymbol{u}^N(\tau_k))$ where the superscript N denotes the N^{th} degree approximation of the relevant variables. The same notation holds for all other variables. Covectors are denoted by column vectors than row vectors to conform with the notion of a gradient as a column vector.

Under suitable constraint qualifications[11], the Minimum Principle isolates possible optimal solutions to Problem B by a search for vector-covector pairs in the primal-dual space. Denoting this as Problem B$^\lambda$, it is defined as:

Problem $\mathbf{B}^{\boldsymbol{\lambda}}$

Determine the state-control-covector function 4-tuple, $[\tau_0, \tau_f] \ni \tau \mapsto \{\boldsymbol{x} \in \mathbb{R}^{N_x}, \boldsymbol{u} \in \mathbb{R}^{N_u}, \boldsymbol{\lambda} \in \mathbb{R}^{N_x}, \boldsymbol{\mu} \in \mathbb{R}_+^{N_h}\}$, a covector $\boldsymbol{\nu} \in \mathbb{R}^{N_e}$, and the clock times τ_0 and τ_f that satisfy Eqs.(2)-(4) in addition to the following conditions:

$$\dot{\boldsymbol{\lambda}}(\tau) = -\frac{\partial L[\tau]}{\partial \boldsymbol{x}} \tag{5}$$

$$\frac{\partial L}{\partial \boldsymbol{u}} = \mathbf{0} \tag{6}$$

$$\{\boldsymbol{\lambda}(\tau_0), \boldsymbol{\lambda}(\tau_f)\} = \left\{-\frac{\partial E_e}{\partial \boldsymbol{x}(\tau_0)}, \frac{\partial E_e}{\partial \boldsymbol{x}(\tau_f)}\right\} \tag{7}$$

$$\{H[\tau_0], H[\tau_f]\} = \left\{\frac{\partial E_e}{\partial \tau_0}, -\frac{\partial E_e}{\partial \tau_f}\right\} \tag{8}$$

where L is the D-form of the Lagrangian of the Hamiltonian defined as[11],

$$L(\boldsymbol{x}, \boldsymbol{u}, \boldsymbol{\lambda}, \boldsymbol{\mu}) = H(\boldsymbol{x}, \boldsymbol{u}, \boldsymbol{\lambda}) + \boldsymbol{\mu}^T \mathbf{h}(\boldsymbol{x}, \boldsymbol{u}) \tag{9}$$

where H is the (unminimized) Hamiltonian,

$$H(\boldsymbol{x}, \boldsymbol{u}, \boldsymbol{\lambda}) = \boldsymbol{\lambda}^T \mathbf{f}(\boldsymbol{x}, \boldsymbol{u}) + F(\boldsymbol{x}, \boldsymbol{u}) \tag{10}$$

and $\boldsymbol{\mu} \in \mathbb{R}_+^{N_h}$ satisfies the complementarity condition,

$$\boldsymbol{\mu}^T(\tau)\mathbf{h}[\tau] = 0 \qquad \forall \tau \in [\tau_0, \tau_f] \tag{11}$$

In the above equations, E_e is defined as

$$E_e(\boldsymbol{x}(\tau_0), \boldsymbol{x}(\tau_f), \tau_0, \tau_f, \boldsymbol{\nu}) = E(\boldsymbol{x}(\tau_f), \boldsymbol{x}(\tau_0), \tau_0, \tau_f) + \boldsymbol{\nu}^T \mathbf{e}(\boldsymbol{x}(\tau_0), \boldsymbol{x}(\tau_f), \tau_0, \tau_f) \tag{12}$$

If the path constraint, Eq.(4), is independent of the control (i.e. a pure state constraint), then the costate, $\boldsymbol{\lambda}(\tau)$, must satisfy the jump condition[11],

$$\boldsymbol{\lambda}^-(\tau_e) = \boldsymbol{\lambda}^+(\tau_e) + \left(\frac{\partial \mathbf{h}}{\partial \boldsymbol{x}(\tau_e)}\right)^T \boldsymbol{\eta} \tag{13}$$

where $\boldsymbol{\eta} \in \mathbb{R}^{N_h}$ is a (constant) covector which effectively arises as a result of the implied interior point constraint (with a pure state constraint),

$$\mathbf{h}(\boldsymbol{x}(\tau_e)) = \mathbf{0} \tag{14}$$

where τ_e denotes the entry or exit point of the trajectory. The important point to note about the jump condition, Eq.(13), is that it is derived by explicitly imposing the constraint, Eq.(14). This is important from a control-theoretic point of view but as will be apparent from the results to follow *in the Legendre pseudospectral method, it is not necessary to explicitly impose this constraint.* In fact, the method automatically determines an approximation to the covector jump as part of the solution.

3 The Legendre Pseudospectral Method

The Legendre pseudospectral method is based on interpolating functions on Legendre-Gauss-Lobatto (LGL) quadrature nodes[4]. These points which are distributed over the interval $[-1, 1]$ are given by $t_0 = -1, \quad t_N = 1$, and for $1 \leq l \leq N-1$, t_l are the zeros of $\dot{\mathsf{L}}_N$, the derivative of the Legendre polynomial of degree N, L_N. Using the affine transformation,

$$\tau(t) = \frac{(\tau_f - \tau_0)t + (\tau_f + \tau_0)}{2} \tag{15}$$

that shifts the LGL nodes from the computational domain $t \in [-1, 1]$ to the physical domain $\tau \in [\tau_0, \tau_f]$, the state and control functions are approximated by Nth degree polynomials of the form

$$\boldsymbol{x}(\tau(t)) \approx \boldsymbol{x}^N(\tau(t)) = \sum_{l=0}^{N} \boldsymbol{x}_l \phi_l(t) \tag{16}$$

$$\boldsymbol{u}(\tau(t)) \approx \boldsymbol{u}^N(\tau(t)) = \sum_{l=0}^{N} \boldsymbol{u}_l \phi_l(t) \tag{17}$$

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

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

are the Lagrange interpolating polynomials of order N. It can be verified that,

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

Hence, it follows that $\boldsymbol{x}_l = \boldsymbol{x}^N(\tau_l)$, $\boldsymbol{u}_l = \boldsymbol{u}^N(\tau_l)$ where $\tau_l = \tau(t_l)$ so that $\tau_N \equiv \tau_f$. Next, differentiating Eq. (16) and evaluating it at the node points, t_k, results in

$$\dot{\boldsymbol{x}}^N(\tau_k) = \frac{d\boldsymbol{x}^N}{d\tau}\Big|_{\tau=\tau_k} = \frac{d\boldsymbol{x}^N}{dt}\frac{dt}{d\tau}\Big|_{t_k} = \frac{2}{\tau_f - \tau_0}\sum_{l=0}^{N} D_{kl}\boldsymbol{x}_l \equiv \frac{2}{\tau_f - \tau_0}\mathbf{d}_k \tag{18}$$

where $D_{kl} = \dot{\phi}_l(t_k)$ are entries of the $(N+1) \times (N+1)$ differentiation matrix $\mathbf{D}$ [4]

$$\mathbf{D} := [D_{kl}] := \begin{cases} \frac{\mathsf{L}_N(t_k)}{\mathsf{L}_N(t_l)} \cdot \frac{1}{t_k - t_l} & k \neq l \\ -\frac{N(N+1)}{4} & k = l = 0 \\ \frac{N(N+1)}{4} & k = l = N \\ 0 & \text{otherwise} \end{cases} \tag{19}$$

This facilitates the approximation of the state dynamics to the following algebraic equations

$$\frac{\tau_f - \tau_0}{2}\mathbf{f}(\boldsymbol{x}_k, \boldsymbol{u}_k) - \sum_{l=0}^{N} D_{kl}\boldsymbol{x}_l = \mathbf{0} \qquad k = 0, \ldots, N$$

Approximating the Bolza cost function, Eq.(1), by the Gauss-Lobatto integration rule, we get,

$$J[\boldsymbol{X}^N, \boldsymbol{U}^N, \tau_0, \tau_f] = E(\boldsymbol{x}_0, \boldsymbol{x}_N, \tau_0, \tau_f) + \frac{\tau_f - \tau_0}{2}\sum_{k=0}^{N} F(\boldsymbol{x}_k, \boldsymbol{u}_k)w_k$$

where

$$\boldsymbol{X}^N = [\boldsymbol{x}_0; \boldsymbol{x}_1; \ldots; \boldsymbol{x}_N], \quad \boldsymbol{U}^N = [\boldsymbol{u}_0; \boldsymbol{u}_1; \ldots; \boldsymbol{u}_N]$$

and w_k are the LGL weights given by

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

Thus, Problem B is discretized by the following nonlinear programming (NLP) problem:

Problem B$^{\boldsymbol{N}}$

Find the $(N+1)(N_x+N_u)+2$ vector $\boldsymbol{X}_{NP} = (\boldsymbol{X}^N; \boldsymbol{U}^N; \tau_0, \tau_f)$ that minimizes

$$J(\boldsymbol{X}_{NP}) \equiv J^N = E(\boldsymbol{x}_0, \boldsymbol{x}_N, \tau_0, \tau_f) + \frac{\tau_f - \tau_0}{2}\sum_{k=0}^{N} F(\boldsymbol{x}_k, \boldsymbol{u}_k)w_k \tag{20}$$

subject to

$$\frac{\tau_f - \tau_0}{2}\mathbf{f}(\boldsymbol{x}_k, \boldsymbol{u}_k) - \sum_{l=0}^{N} D_{kl}\boldsymbol{x}_l = \mathbf{0} \tag{21}$$

$$\mathbf{e}(\boldsymbol{x}_0, \boldsymbol{x}_N, \tau_0, \tau_f) = \mathbf{0} \tag{22}$$

$$\mathbf{h}(\boldsymbol{x}_k, \boldsymbol{u}_k) \leq \mathbf{0} \tag{23}$$

for $k = 0, \ldots, N$.

Problem B^λ can also be discretized in much the same manner. Approximating the costate by the N^{th} degree polynomial,

$$\boldsymbol{\lambda}(\tau(t)) \approx \boldsymbol{\lambda}^N(\tau(t)) = \sum_{l=0}^{N} \boldsymbol{\lambda}_l \phi_l(t) \tag{24}$$

and letting $\boldsymbol{\Lambda}_{NP} = [\boldsymbol{\lambda}_0; \boldsymbol{\lambda}_1; \ldots; \boldsymbol{\lambda}_N; \boldsymbol{\mu}_0; \boldsymbol{\mu}_1; \ldots; \boldsymbol{\mu}_N; \boldsymbol{\nu}_0; \boldsymbol{\nu}_f]$, we can discretize Problem B^λ as,

Problem $\mathbf{B}^{\boldsymbol{\lambda} \boldsymbol{N}}$

Find $\boldsymbol{X}_{NP}$ and $\boldsymbol{\Lambda}_{NP}$ that satisfy Eqs.(21)-(23) in addition to the following nonlinear algebraic relations:

$$\sum_{l=0}^{N} D_{kl}\boldsymbol{\lambda}_l = -\frac{\partial L_k}{\partial \boldsymbol{x}_k} \tag{25}$$

$$\frac{\partial L_k}{\partial \boldsymbol{u}_k} = \mathbf{0} \tag{26}$$

$$\{\boldsymbol{\lambda}_0, \boldsymbol{\lambda}_N\} = \left\{-\frac{\partial E_e}{\partial \boldsymbol{x}_0}, \frac{\partial E_e}{\partial \boldsymbol{x}_N}\right\} \tag{27}$$

$$\{H_0, H_N\} = \left\{\frac{\partial E_e}{\partial \tau_0}, -\frac{\partial E_e}{\partial \tau_N}\right\} \tag{28}$$

$$\boldsymbol{\mu}_k^T \mathbf{h}_k = 0, \quad \boldsymbol{\mu}_k \geq \mathbf{0} \tag{29}$$

for $k = 0, \ldots, N$.

Remark 1. In the case of pure state constraints, it is necessary to determine *a priori* a switching structure and impose the jump conditions for optimality. Assuming a sufficiently large N, the jump condition can be approximated as,

$$\boldsymbol{\lambda}(t_e) = \boldsymbol{\lambda}(t_{e+1}) + \left(\frac{\partial \mathbf{h}(\boldsymbol{x}_e)}{\partial \boldsymbol{x}_e}\right)^T \boldsymbol{\eta} \tag{30}$$

for all points t_e that are the junction points of the switching structure. This is the indirect Legendre pseudospectral method[8] and represents a discretization of the multi-point boundary value problem. It is obvious that the direct method (Problem B^N) is far simpler to implement than the indirect method. This is true of any direct/indirect method[2]. However, unlike the indirect method, not much can be said about the optimality or the convergence of the direct method. The theorem of the next section shows how to get the high performance of the indirect method without actually implementing it by way of the significantly simpler implementation of the direct method.

3.1 KKT Conditions for Problem $\mathbf{B}^{\boldsymbol{N}}$

The Lagrangian for Problem B^N, can be written as

$$\overline{J}^N(\boldsymbol{X}_{NP}, \widetilde{\boldsymbol{\nu}}, \widetilde{\boldsymbol{\lambda}}, \widetilde{\boldsymbol{\mu}}) = J^N(\boldsymbol{X}_{NP}) + \widetilde{\boldsymbol{\nu}}^T \mathbf{e}(\boldsymbol{x}_0, \boldsymbol{x}_N, \tau_0, \tau_f) +$$
$$\sum_{i=0}^{N} \left(\widetilde{\boldsymbol{\lambda}}_i^T \{(\frac{\tau_f - \tau_0}{2})\mathbf{f}_i(\boldsymbol{X}_{NP}) - \mathbf{d}_i(\boldsymbol{X}^N)\} + \widetilde{\boldsymbol{\mu}}_i^T \mathbf{h}_i(\boldsymbol{X}_{NP})\right) \tag{31}$$

where $\widetilde{\boldsymbol{\nu}}, \widetilde{\boldsymbol{\lambda}}_i, \widetilde{\boldsymbol{\mu}}_i$ are the KKT multipliers associated with the NLP. Using Lemma 1 below, the KKT conditions may be written quite succinctly in a certain form described later in this section.

Lemma 1. *The elements of the Differentiation Matrix, D_{ik}, and the LGL weights, w_i, together satisfy the following properties,*

$$w_i D_{ik} + w_k D_{ki} = 0 \qquad i, k = 1, \ldots, N-1 \tag{32}$$

For the boundary terms, we have $2w_0 D_{00} = -1$, and $2w_N D_{NN} = 1$. Further, $\sum_{i=0}^{N} w_i = 2$.

For a proof of this, please see [9].

Lemma 2. *The LGL-weight-normalized multipliers $\frac{\widetilde{\boldsymbol{\lambda}}_k}{w_k}, \frac{\widetilde{\boldsymbol{\mu}}_k}{w_k}$ satisfy the same equations as the discrete costates (Cf. Eq.(25)) at the interior nodes, $k = 1, \ldots N-1$; i.e., we have*

$$\frac{\partial L}{\partial \boldsymbol{x}_k}(\boldsymbol{x}_k, \boldsymbol{u}_k, \frac{\widetilde{\boldsymbol{\lambda}}_k}{w_k}, \frac{\widetilde{\boldsymbol{\mu}}_k}{w_k}) + \sum_{i=0}^{N} D_{ki} \left(\frac{\widetilde{\boldsymbol{\lambda}}_i}{w_i} \right) = \mathbf{0} \tag{33}$$

Proof: Consider the interior state variables $(\boldsymbol{x}_1, \ldots \boldsymbol{x}_{N-1})$. From applying the KKT condition at the interior nodes to Eq.(31), i.e. $\frac{\partial \bar{J}^N}{\partial \boldsymbol{x}_k} = \mathbf{0}$, we have

$$\frac{\partial}{\partial \boldsymbol{x}_k} \left[\sum_{i=0}^{N} \widetilde{\boldsymbol{\lambda}}_i^T \left[\frac{\tau_f - \tau_0}{2} \mathbf{f}_i - \mathbf{d}_i \right] + \widetilde{\boldsymbol{\mu}}_i^T \mathbf{h}_i \right] = -\frac{\partial J^N}{\partial \boldsymbol{x}_k} \tag{34}$$

Since the functions $\mathbf{f}, \mathbf{h}, F$ are evaluated only at the points t_i, we have

$$\frac{\partial}{\partial \mathbf{x}_k} \left[\sum_{i=0}^{N} \widetilde{\boldsymbol{\lambda}}_i^T \left(\frac{\tau_f - \tau_0}{2} \mathbf{f}_i \right) + \widetilde{\boldsymbol{\mu}}_i^T \mathbf{h}_i + \frac{\tau_f - \tau_0}{2} F_i w_i \right] = \frac{\tau_f - \tau_0}{2} \left(\frac{\partial \mathbf{f}_k}{\partial \boldsymbol{x}_k} \right)^T \widetilde{\boldsymbol{\lambda}}_k + \frac{\tau_f - \tau_0}{2} \frac{\partial F_k}{\partial \boldsymbol{x}_k} w_k + \left(\frac{\partial \mathbf{h}_k}{\partial \boldsymbol{x}_k} \right)^T \widetilde{\boldsymbol{\mu}}_k \tag{35}$$

For the term involving the state derivatives, a more complicated expression is obtained since the differentiation matrix $\mathbf{D}$ relates the different components of $\boldsymbol{x}_k$:

$$\frac{\partial}{\partial \boldsymbol{x}_k} \left[\sum_{i=0}^{N} \widetilde{\boldsymbol{\lambda}}_i^T \mathbf{d}_i \right] = \sum_{i=0}^{N} D_{ik} \widetilde{\boldsymbol{\lambda}}_i \tag{36}$$

From Lemma 1, $D_{ik} = -\frac{w_k}{w_i} D_{ki}$, therefore by putting together Eqs. (35)-(36), the following is obtained for $k = 1, \ldots, N-1$:

$$\frac{\tau_f - \tau_0}{2} \frac{\partial F_k}{\partial \boldsymbol{x}_k} w_k + \frac{\tau_f - \tau_0}{2} \left(\frac{\partial \mathbf{f}_k}{\partial \boldsymbol{x}_k} \right)^T \widetilde{\boldsymbol{\lambda}}_k + w_k \sum_{i=0}^{N} D_{ki} \left(\frac{\widetilde{\boldsymbol{\lambda}}_i}{w_i} \right) + \left(\frac{\partial \mathbf{h}_k}{\partial \boldsymbol{x}_k} \right)^T \widetilde{\boldsymbol{\mu}}_k = \mathbf{0} \tag{37}$$

Dividing Eq. (37) by w_k yields the desired result for $k = 1, \ldots, N-1$. □

Lemma 3. *The LGL-weight-normalized multipliers* $\frac{\widetilde{\boldsymbol{\lambda}}_k}{w_k}, \frac{\widetilde{\boldsymbol{\mu}}_k}{w_k}$ *satisfy the discrete first-order optimality condition associated with the minimization of the Hamiltonian at all node points:*

$$\frac{\partial L}{\partial \boldsymbol{u}_k}(\boldsymbol{x}_k, \boldsymbol{u}_k, \frac{\widetilde{\boldsymbol{\lambda}}_k}{w_k}, \frac{\widetilde{\boldsymbol{\mu}}_k}{w_k}) = \mathbf{0} \tag{38}$$

Proof: Considering the terms that involve differentiation with respect to the control variables $\boldsymbol{u}_k$ in Eq. (31) yields

$$\left(\frac{\tau_f - \tau_0}{2}\frac{\partial \mathbf{f}_k}{\partial \boldsymbol{u}_k}\right)^T \widetilde{\boldsymbol{\lambda}}_k + \left(\frac{\partial \mathbf{h}_k}{\partial \boldsymbol{u}_k}\right)^T \widetilde{\boldsymbol{\mu}}_k = -\frac{\partial J^N}{\partial \boldsymbol{u}_k} \quad k = 0, \ldots, N. \tag{39}$$

Since

$$\frac{\partial J^N}{\partial \boldsymbol{u}_k} = \left(\frac{\tau_f - \tau_0}{2}\right)\frac{\partial F_k}{\partial \boldsymbol{u}_k} w_k \tag{40}$$

dividing Eq.(39) by w_k, yields the desired result. □

Lemma 4. *At the final node, the KKT multipliers satisfy the following equation:*

$$w_N \left(\frac{\partial L}{\partial \boldsymbol{x}_N}(\boldsymbol{x}_N, \boldsymbol{u}_N, \frac{\widetilde{\boldsymbol{\lambda}}_N}{w_N}, \frac{\widetilde{\boldsymbol{\mu}}_N}{w_N}) + \sum_{i=0}^{N} D_{Ni}\frac{\widetilde{\boldsymbol{\lambda}}_i}{w_i}\right) \equiv \mathbf{c}_N \tag{41}$$

$$\frac{\widetilde{\boldsymbol{\lambda}}_N}{w_N} - \frac{\partial \tilde{E}_e}{\partial \boldsymbol{x}_N} \equiv \mathbf{c}_N \tag{42}$$

where $\tilde{E}_e = E_e(\boldsymbol{x}_0, \boldsymbol{x}_N, \tau_0, \tau_N, \widetilde{\boldsymbol{\nu}})$

Proof: The following KKT condition holds for the last node:

$$\left(\frac{\partial \mathbf{e}}{\partial \boldsymbol{x}_N}\right)^T \widetilde{\boldsymbol{\nu}} + \frac{\tau_f - \tau_0}{2}\left(\frac{\partial \mathbf{f}_N}{\partial \boldsymbol{x}_N}\right)^T \widetilde{\boldsymbol{\lambda}}_N - \sum_{i=0}^{N} D_{iN}\widetilde{\boldsymbol{\lambda}}_i + \left(\frac{\partial \mathbf{h}_N}{\partial \boldsymbol{x}_N}\right)^T \widetilde{\boldsymbol{\mu}}_N = -\frac{\partial J^N}{\partial \boldsymbol{x}_N} \tag{43}$$

Using the relationship

$$D_{iN} = -\frac{w_N}{w_i} D_{Ni}, \quad i \neq N \quad \text{and} \quad 2D_{NN} = \frac{1}{w_N}$$

and adding $2D_{NN}\widetilde{\boldsymbol{\lambda}}_N = \frac{\widetilde{\boldsymbol{\lambda}}_N}{w_N}$ to both sides of Eqn (43) and rearranging the terms, the following is obtained:

$$\left(\frac{\tau_f - \tau_0}{2}\left(\frac{\partial F_N}{\partial \boldsymbol{x}_N}\right) w_N + \frac{\tau_f - \tau_0}{2}\left(\frac{\partial \mathbf{f}_N}{\partial \boldsymbol{x}_N}\right)^T \widetilde{\boldsymbol{\lambda}}_N + w_N \sum_{i=0}^{N} D_{Ni}\frac{\widetilde{\boldsymbol{\lambda}}_i}{w_i} + \left(\frac{\partial \mathbf{h}_N}{\partial \boldsymbol{x}_N}\right)^T \widetilde{\boldsymbol{\mu}}_N\right) =$$

$$2D_{NN}\widetilde{\boldsymbol{\lambda}}_N - \frac{\partial E}{\partial \boldsymbol{x}_N} - \left(\frac{\partial \mathbf{e}}{\partial \boldsymbol{x}_N}\right)^T \widetilde{\boldsymbol{\nu}} \tag{44}$$

or

$$w_N\left(\frac{\partial L}{\partial \boldsymbol{x}_N}(\boldsymbol{x}_N,\boldsymbol{u}_N,\frac{\widetilde{\boldsymbol{\lambda}}_N}{w_N},\frac{\widetilde{\boldsymbol{\mu}}_N}{w_N}) + \sum_{i=0}^{N} D_{Ni}\frac{\widetilde{\boldsymbol{\lambda}}_i}{w_i}\right) = \frac{\widetilde{\boldsymbol{\lambda}}_N}{w_N} - \frac{\partial \tilde{E}_e}{\partial \boldsymbol{x}_N} \equiv \mathbf{c}_N. \ \Box$$

Corollary 1. *The result for the zeroth node (i.e. initial time condition) can be shown in a similar fashion:*

$$-w_0\left(\frac{\partial L}{\partial \boldsymbol{x}_0}(\boldsymbol{x}_0,\boldsymbol{u}_0,\frac{\widetilde{\boldsymbol{\lambda}}_0}{w_0},\frac{\widetilde{\boldsymbol{\mu}}_0}{w_0}) + \sum_{i=0}^{N} D_{0i}\frac{\widetilde{\boldsymbol{\lambda}}_i}{w_i}\right) = \frac{\widetilde{\boldsymbol{\lambda}}_0}{w_0} + \frac{\partial \tilde{E}_e}{\partial \boldsymbol{x}_0} \equiv \mathbf{c}_0$$

Lemma 5. *The Lagrange multipliers $\widetilde{\boldsymbol{\lambda}}_i$ and $\widetilde{\boldsymbol{\nu}}$ satisfy the condition,*

$$\frac{1}{2}\sum_{i=0}^{N} w_i H\big(\boldsymbol{x}_i,\boldsymbol{u}_i,\frac{\widetilde{\boldsymbol{\lambda}}_i}{w_i}\big) = -\frac{\partial \tilde{E}_e}{\partial \tau_N} \tag{45}$$

$$\frac{1}{2}\sum_{i=0}^{N} w_i H\big(\boldsymbol{x}_i,\boldsymbol{u}_i,\frac{\widetilde{\boldsymbol{\lambda}}_i}{w_i}\big) = \frac{\partial \tilde{E}_e}{\partial \tau_0} \tag{46}$$

Proof: Applying the KKT condition for the variable, τ_N, we have,

$$-\frac{\partial E}{\partial \tau_N} - \frac{\partial \mathbf{e}}{\partial \tau_N}^T \widetilde{\boldsymbol{\nu}} = \left[\sum_{i=0}^{N} \frac{\widetilde{\boldsymbol{\lambda}}_i^T \mathbf{f}_i}{2} + \frac{F_i w_i}{2}\right] = \frac{1}{2}\sum_{i=0}^{N} w_i\left(F_i + \frac{\widetilde{\boldsymbol{\lambda}}_i^T}{w_i}\mathbf{f}_i\right)$$

and hence the first part of the lemma. The second part of the lemma follows similarly by considering the variable τ_0. $\Box$

Collecting all these results, and letting

$$\widetilde{\boldsymbol{\Lambda}}_{NP} = [\widetilde{\boldsymbol{\lambda}}_0;\widetilde{\boldsymbol{\lambda}}_1;\ldots;\widetilde{\boldsymbol{\lambda}}_N;\widetilde{\boldsymbol{\mu}}_0;\widetilde{\boldsymbol{\mu}}_1;\ldots;\widetilde{\boldsymbol{\mu}}_N;\widetilde{\boldsymbol{\nu}}_0;\widetilde{\boldsymbol{\nu}}_f]$$

the dualization of Problem B^N may be cast in terms of Problem $\mathrm{B}^{N\lambda}$:

Problem $\mathrm{B}^{N\boldsymbol{\lambda}}$

Find $\boldsymbol{X}_{NP}$ and $\widetilde{\boldsymbol{\Lambda}}_{NP}$ that satisfy Eqs.(21)-(23) in addition to the following nonlinear algebraic relations:

$$\frac{\partial L}{\partial \boldsymbol{u}_k}(\boldsymbol{x}_k,\boldsymbol{u}_k,\frac{\widetilde{\boldsymbol{\lambda}}_k}{w_k},\frac{\widetilde{\boldsymbol{\mu}}_k}{w_k}) = \mathbf{0} \qquad k=0,\ldots,N \tag{47}$$

$$\widetilde{\boldsymbol{\mu}}_k^T\mathbf{h}_k = 0, \quad \widetilde{\boldsymbol{\mu}}_k \geq \mathbf{0} \qquad k=0,\ldots,N \tag{48}$$

$$\frac{\partial L}{\partial \boldsymbol{x}_k}(\boldsymbol{x}_k,\boldsymbol{u}_k,\frac{\widetilde{\boldsymbol{\lambda}}_k}{w_k},\frac{\widetilde{\boldsymbol{\mu}}_k}{w_k}) + \sum_{i=0}^{N} D_{ki}\left(\frac{\widetilde{\boldsymbol{\lambda}}_i}{w_i}\right) = \mathbf{0} \qquad k=1,\ldots,N-1 \tag{49}$$

and

$$\frac{\partial L}{\partial \boldsymbol{x}_N}(\boldsymbol{x}_N, \boldsymbol{u}_N, \frac{\widetilde{\boldsymbol{\lambda}}_N}{w_N}, \frac{\widetilde{\boldsymbol{\mu}}_N}{w_N}) + \sum_{i=0}^{N} D_{Ni}\frac{\widetilde{\boldsymbol{\lambda}}_i}{w_i} = \frac{\mathbf{c}_N}{w_N} \tag{50}$$

$$\frac{\widetilde{\boldsymbol{\lambda}}_N}{w_N} - \frac{\partial \tilde{E}_e}{\partial \boldsymbol{x}_N} = \mathbf{c}_N \tag{51}$$

$$\frac{\partial L}{\partial \boldsymbol{x}_0}(\boldsymbol{x}_0, \boldsymbol{u}_0, \frac{\widetilde{\boldsymbol{\lambda}}_0}{w_0}, \frac{\widetilde{\boldsymbol{\mu}}_0}{w_0}) + \sum_{i=0}^{N} D_{0i}\frac{\widetilde{\boldsymbol{\lambda}}_i}{w_i} = -\frac{\mathbf{c}_0}{w_0} \tag{52}$$

$$\frac{\widetilde{\boldsymbol{\lambda}}_0}{w_0} + \frac{\partial \tilde{E}_e}{\partial \boldsymbol{x}_0} = \mathbf{c}_0 \tag{53}$$

$$\frac{1}{2}\sum_{i=0}^{N} w_i H\big(\boldsymbol{x}_i, \boldsymbol{u}_i, \frac{\widetilde{\boldsymbol{\lambda}}_i}{w_i}\big) = -\frac{\partial \tilde{E}_e}{\partial \tau_N} \tag{54}$$

$$\frac{1}{2}\sum_{i=0}^{N} w_i H\big(\boldsymbol{x}_i, \boldsymbol{u}_i, \frac{\widetilde{\boldsymbol{\lambda}}_i}{w_i}\big) = \frac{\partial \tilde{E}_e}{\partial \tau_0} \tag{55}$$

where $\mathbf{c}_0$ and $\mathbf{c}_N$ are arbitrary vectors in $\mathbb{R}^{N_x}$. The deliberate formulation of the KKT conditions for Problem B^N in the above form facilitates a definition of Closure Conditions:

Definition 1. *Closure Conditions are defined as the set of constraints that must be added to Problem $B^{N\lambda}$ so that every solution of this restricted problem is equivalent to the solution of Problem $B^{\lambda N}$*

From this definition, the Closure Conditions are obtained by simply matching the equations for Problems $\mathrm{B}^{N\lambda}$ to those of Problem $\mathrm{B}^{\lambda N}$. This results in,

$$\mathbf{c}_0 = \mathbf{0} \tag{56}$$

$$\mathbf{c}_N = \mathbf{0} \tag{57}$$

$$\frac{1}{2}\sum_{i=0}^{N} w_i H\big(\boldsymbol{x}_i, \boldsymbol{u}_i, \frac{\widetilde{\boldsymbol{\lambda}}_i}{w_i}\big) = H_0 = H_N \tag{58}$$

The Closure Conditions facilitate our main theorem:

The Covector Mapping Theorem

Theorem 1. *There exist Lagrange multipliers $\widetilde{\boldsymbol{\lambda}}_i, \widetilde{\boldsymbol{\mu}}_i$ that are equal to the pseudospectral approximations of the covectors $\boldsymbol{\lambda}^N(\tau_i), \boldsymbol{\mu}^N(\tau_i)$ at the shifted LGL node τ_i multiplied by the corresponding LGL weight w_i. Further, there exists a $\widetilde{\boldsymbol{\nu}}$ that is equal to the constant covector $\boldsymbol{\nu}$. In other words, we can write,*

$$\boldsymbol{\lambda}^N(\tau_i) = \frac{\widetilde{\boldsymbol{\lambda}}_i}{w_i}, \quad \boldsymbol{\mu}^N(\tau_i) = \frac{\widetilde{\boldsymbol{\mu}}_i}{w_i}, \quad \widetilde{\boldsymbol{\nu}} = \boldsymbol{\nu} \tag{59}$$

3.2 Proof of the Theorem

Since a solution, $\{\boldsymbol{x}_i, \boldsymbol{u}_i, \boldsymbol{\lambda}_i, \boldsymbol{\mu}_i, \boldsymbol{\nu}\}$, to Problem $\mathrm{B}^{\lambda N}$ exists (by assumption), it follows that $\{\boldsymbol{x}_i, \boldsymbol{u}_i, w_i\boldsymbol{\lambda}_i, w_i\boldsymbol{\mu}_i, \boldsymbol{\nu}\}$ solves Problem $\mathrm{B}^{N\lambda}$ while automatically satisfying the Closure Conditions. Conversely, a solution, $\{\boldsymbol{x}_i, \boldsymbol{u}_i, \widetilde{\boldsymbol{\lambda}}_i, \widetilde{\boldsymbol{\mu}}_i, \widetilde{\boldsymbol{\nu}}\}$, of Problem $\mathrm{B}^{N\lambda}$ that satisfies the Closure Conditions provides a solution, $\{\boldsymbol{x}_i, \boldsymbol{u}_i, \frac{\widetilde{\boldsymbol{\lambda}}_i}{w_i}, \frac{\widetilde{\boldsymbol{\mu}}_i}{w_i}, \widetilde{\boldsymbol{\nu}}\}$, to Problem $\mathrm{B}^{\lambda N}$. □

Remark 2. A solution of Problem $\mathrm{B}^{\lambda N}$ always provides a solution to Problem $\mathrm{B}^{N\lambda}$; however, the converse is not true in the absence of the Closure Conditions. Thus, the Closure Conditions guarantee an order-preserving bijective map between the solutions of Problem $\mathrm{B}^{N\lambda}$ and $\mathrm{B}^{\lambda N}$. The commutative diagram depicted in Fig.1 captures the core ideas.

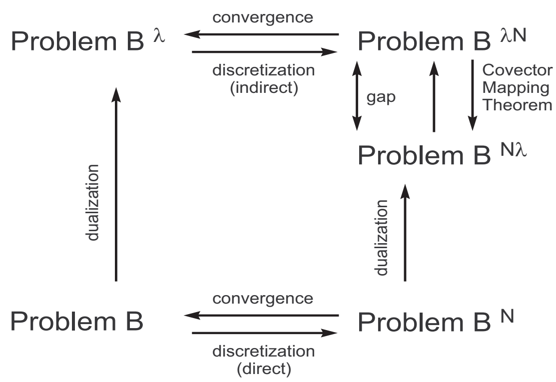

Fig. 1. Commutative Diagram for Discretization and Dualization

Remark 3. The Closure Conditions given by $\mathbf{c}_0 = \mathbf{0} = \mathbf{c}_N$ are a simple requirement of the fact that the PS transformed discrete adjoint equations be satisfied at the end points in addition to meeting the endpoint transversality conditions. On the other hand, the condition given by Eq.(58) states the constancy of the discrete Hamiltonian in a weak form (see Lemma 1).

Remark 4. The Closure Conditions signify the closing of the gap between Problems $\mathrm{B}^{N\lambda}$ and $\mathrm{B}^{\lambda N}$ which exist for any given degree of approximation, N. The issue of convergence of Problem B^N to Problem B via Problem $B^{\lambda N}$ is discussed in Ref.[13].

4 Numerical Example

To illustrate the theory presented in the previous sections, the Breakwell problem[3] is considered:

Minimize

$$J = \frac{1}{2}\int_0^1 u^2\,dt$$

subject to the equations of motion

$$\dot{x}(\tau) = v(\tau), \qquad \dot{v}(\tau) = u(\tau)$$

the boundary conditions

$$x(0) = 0, \quad x(1) = 0, \quad v(0) = 1.0, \quad v(1) = -1.0$$

and the state constraint

$$x(\tau) \leq \ell = 0.1$$

Figures 2 and 3 demonstrate the excellent agreement between the analytical solution[3] and the solution obtained from our Legendre pseudospectral method. The solution was obtained for 50 LGL points with the aid of DIDO[16], a software package that implements our ideas. The cost function obtained is 4.4446 which agrees very well with the analytic optimal result of $J = \frac{4}{9\ell} = 4.4444$. It is apparent that the optimal switching structure is *free-constrained-free*. The costates corresponding to the D-form of the Lagrangian are shown in Figure 4. Note that the method adequately captures the fact that λ_v should be continuous while λ_x should have jump discontinuities given by,[3]

$$\lambda_x^-(\tau_j) - \lambda_x^+(\tau_j) = \frac{2}{9\ell^2} \quad j = 1, 2 \quad \tau_1 = 3\ell, \quad \tau_2 = 1 - 3\ell$$

Figure 4 exhibits a jump discontinuity of 22.2189 which compares very well with the analytical value of 22.2222.

5 Conclusions

A Legendre pseudospectral approximation of the constrained Bolza problem has revealed that there is a loss of information when a dualization is performed after discretization. This information loss can be restored by way of Closure Conditions introduced in this paper. These conditions also facilitate a spectrally accurate way of representing the covectors associated with the continuous problem by way of the Covector Mapping Theorem (CMT). All these results can be succinctly represented by a commutative diagram. The practical advantage of the CMT is that nonlinear optimal control problems can be solved efficiently and accurately without developing the necessary conditions. On the other hand, the optimality of the solution can be checked by using the numerical approximations of the covectors obtained from the CMT. Since these solutions can presently be obtained in a matter of seconds, it appears that the proposed technique can be used for optimal feedback control in the context of a nonlinear model predictive framework.

[3] Ignoring the typographical errors, the costates given in Ref.[3] correspond to the P-form[11] and exhibit a jump discontinuity in λ_v as well.

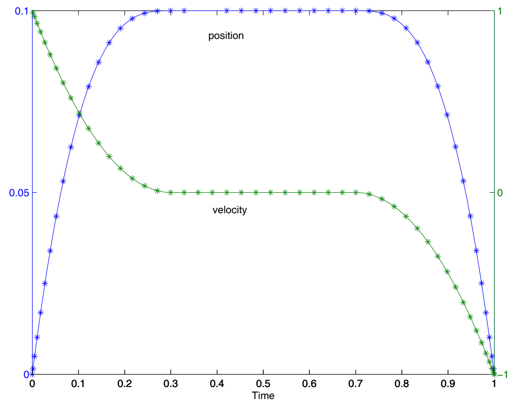

Fig. 2. PS states, x and v. Solid line is analytical.

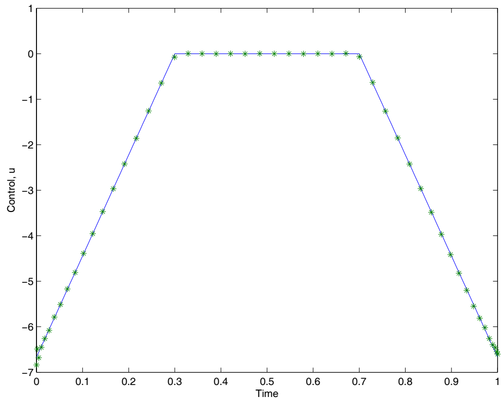

Fig. 3. PS control, u. Solid line is analytical.

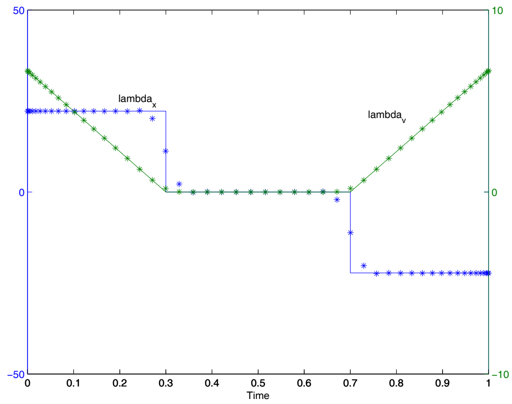

Fig. 4. Costates, λ_x and λ_v from CMT. Solid line is analytical.

References

1. Bellman, R. E. (1957). *Dynamic Programming.* Princeton University Press, Princeton, NJ.
2. Betts, J. T. (1998). Survey of numerical methods for trajectory optimization. *Journal of Guidance, Control, and Dynamics,* Vol 21, No. 2, 193-207.
3. Bryson, A.E., and Ho, Y.C., (1975). *Applied Optimal Control.* Hemisphere, New York.
4. Canuto, C., Hussaini, M. Y., Quarteroni, A., and Zang, T.A. (1988). *Spectral Methods in Fluid Dynamics.* Springer Verlag, New York.
5. Clarke, F. H., Ledyaev, Yu. S., Stern, R. J. and Wolenski, P. R. (1998). *Nonsmooth Analysis and Control Theory,* Springer-Verlag, New York.
6. Elnagar, J., Kazemi, M. A. and Razzaghi, M. (1995). The pseudospectral Legendre method for discretizing optimal control problems. *IEEE Transactions on Automatic Control,* Vol. 40, No. 10, 1793–1796.
7. Elnagar, J. and Razzaghi, M. (1997). A collocation-type method for linear quadratic optimal control problems. *Optimal Control Applications and Methods,* Vol. 18, 227–235.
8. Fahroo, F. and Ross, I. M. (2000). Trajectory optimization by indirect spectral collocation methods. *Proc. AIAA/AAS Astrodynamics Specialist Conference.* Denver, CO, 123–129.
9. Fahroo, F., and Ross, I. M. (2001). Costate estimation by a Legendre pseudospectral method. *Journal of Guidance, Control, and Dynamics*, Vol. 24, No. 2, 270–277.
10. Hager, W. W. (2000). Runge-Kutta methods in optimal control and the transformed adjoint system. *Numerische Mathematik,* Vol. 87, pp. 247-282.
11. Hartl, R. F., Sethi, S. P. and Vickson, R. G. (1995). A survey of the maximum principles for optimal control problems with state constraints. *SIAM Review,* Vol. 37, No. 2, 181–218.
12. Mordukhovich, B. S. and Shvartsman, I. (2002). The approximate maximum principle for constrained control systems. *Proc. 41st IEEE Conf. on Decision an Control*, Las Vegas, NV.
13. Ross, I. M., and Fahroo, F. (2001). Convergence of pseudospectral discretizations of optimal control problems. *Proc. 40th IEEE Conf. on Decision and Control,* Orlando, FL.
14. Ross, I. M. and Fahroo, F. (2002). A perspective on methods for trajectory optimization," *Proc. AIAA/AAS Astrodynamics Specialist Conference*, Monterey, CA, Invited Paper, AIAA-2002-4727.
15. Ross, I. M., and Fahroo, F. (2002). Pseudospectral methods for optimal motion planning of differentially flat systems, *Proc. 41st IEEE Conf. on Decision and Control,* Las Vegas, NV.
16. Ross, I. M. and Fahroo, F., (2002). User's manual for DIDO 2002: A MATLAB application package for dynamic optimization," *NPS Technical Report AA-02-002*, Department of Aeronautics and Astronautics, Naval Postgraduate School, Monterey, CA.
17. Trefethen, L. N., (2000). *Spectral Methods in MATLAB,* SIAM, Philadelphia, PA.