Advances in Pseudospectral Methods for Optimal Control

Fariba Fahroo[*] and I. Michael Ross[†]

Recently, the Legendre pseudospectral (PS) method migrated from theory to flight application onboard the International Space Station for performing a finite-horizon, zero-propellant maneuver. A small technical modification to the Legendre PS method is necessary to manage the limiting conditions at infinity for infinite-horizon optimal control problems. Motivated by these technicalities, the concept of primal-dual weighted interpolation, introduced earlier by the authors, is used to articulate a unified theory for all PS methods for optimal control. This theory illuminates the previously hidden fact of the unit weight function implicit in the the Legendre PS method based on Legendre-Gauss-Lobatto points. The unified framework also reveals why this Legendre PS method is the most appropriate method for solving finite-horizon optimal control problems with arbitrary boundary conditions. This conclusion is borne by a proper definition of orthogonality needed to generate convergent approximations in Hilbert spaces. Special boundary conditions are needed to ensure the convergence of the Legendre PS method based on the Legendre-Gauss-Radau (LGR) and the Legendre-Gauss (LG) points. These facts are illustrated by simple examples and counter examples which reveal when and why PS methods based on LGR and LG points fail. A new kind of consistency in the primal-dual weight functions allows us to generate dual maps (such as Hamiltonians, adjoints etc) without resorting to solving difficult two-point boundary-value problems. These concepts are encapsulated in a unified Covector Mapping Theorem.

I.   Introduction

In 2007, a front page article in *SIAM News*[1] announced that pseudospectral optimal control was successfully used to maneuver the International Space Station. Over the last decade, PS methods have moved rapidly from theory and computation to ground experiments[2,3] ... and now successful flight applications.[1,4,5] The rapid rise of PS methods is, in part, due to its independently reproducible superior performance when compared to other general purpose techniques.[7–10] Furthermore, the versatility of PS methods is evident from the vast number of diverse optimal control problems that have been solved by various practitioners; examples include space station attitude control,[4,5] ascent guidance,[7] interplanetary solar sail mission design,[8] supersonic intercept,[9,10] low-thrust Earth-to-Jupiter rendezvous,[10] loitering of unmanned aerial vehicles,[11,12] lunar guidance,[13] libration-point stationkeeping,[14] momentum-dumping,[15] launch vehicle trajectory optimization,[16] impulsive orbit transfer optimization,[17] inert and electrodynamic tether control,[18,19] magnetic attitude control[20] and many more.[21] Solutions to well over a hundred problems are documented in Ref. [21]. Many of the preceding example problems have been solved by OTIS[22] and DIDO,[23] the FORTRAN and MATLAB codes respectively for solving optimal control problems.

The most widely used PS method is the Legendre PS method[24–26] that is based on the Legendre-Gauss-Lobatto (LGL) node points. This is simply because a rich number of convergence theorems have been proven in the literature,[28–30] and hence one may use this method with a high degree of confidence in its validity. In addition, in practical applications, one may treat the outcome of the LGL/PS method in terms of a rigorous application of the Pontryagin Minimum Principle and accept or reject solutions based on the optimality conditions.[31] This comfort of guarantees is enunciated as the Covector Mapping Principle[27,28,32–35] (CMP) which provides the foundations for generating explicit Covector Mapping Theorems.[25,26,28,29] The proof

---

[*]Professor, Department of Applied Mathematics, Naval Postgraduate School, Monterey, CA 93943. Email: ffahroo@nps.edu. Associate Fellow, AIAA.

[†]Professor and Program Director, Control & Optimization, Naval Postgraduate School, Monterey, CA 93943; im-ross@nps.edu. Associate Fellow, AIAA. http://www.nps.edu/academics/gnclab/

of this theorem utilizes specific quadrature formulas[36, 37] that have thus far shown to be valid only for the Legendre PS method, and hence the popularity of this approach.

In 2005, we proposed[38] a PS method based on Legendre-Gauss-Radau (LGR) points to solve infinite-horizon optimal control problems. This proposal was motivated by a means to handle the singularity problems that arise as a result of transforming the infinite-time domain to a finite time-domain. Additional studies showed[39] a number of other advantages of the shifted LGR/PS method, particularly for real-time applications. This generates a natural question: does the LGR/PS method satisfy CMP-consistency? An apparently simple way to investigate this issue is to apply the CMP to derive an explicit covector map[34] in a manner similar to that of the LGL/PS approach. While conceptually simple, this task is not altogether straightforward as a key lemma related to an integration-by-parts formula[26] is not readily available for non-LGL methods. This crucial formula identifies the correct finite-dimensional inner-product space that is necessary for the construction of the discretized 1-form that defines the sequences of discretized Lagrangians that converge to the continuous-time Lagrangian.[40] In Ref. [41], we showed that a special weight function met all the appropriate criteria for the CMP leading to a new theorem on dual consistency for the LGR/PS method.

Motivated by these ideas for the LGR/PS method and their possible connections to the LGL/PS method, we develop a unified approach to investigate all PS methods. *The key idea introduced in this paper is the notion of weighed interpolants, their duals, and their direct effect on the generation of the correct pre-Hilbert space where the computed solutions lie.* The concept of a pre-Hilbert space is crucial for proofs of convergence theorems[28–30] that rely on the separability of the infinite-dimensional Hilbert spaces used to construct highly-accurate solutions to practical optimal control problems. We presented some of these ideas in Ref. [21]; here, we substantially expand and clarify our previous results.[21]

We begin this paper at the level of first principles by providing a unified set of foundations for all PS methods for optimal control; i.e. PS methods over an arbitrary grid. In limiting the scope of the discussions, we focus on three Legendre PS methods: LGL, LGR and Legendre-Gauss (LG). From the perspective of unification, it is shown that the LGL/PS method is indeed the correct Legendre PS method for solving boundary value problems (BVPs) with arbitrary boundary conditions and hence all generic finite-horizon optimal control problems. The unification also illustrates why the LGR/PS method is indeed better than the LGL/PS method for stabilizing control systems as this problem falls under the realm of a special optimal control problem, the specialty being the infiniteness of the horizon. We illustrate some key consequences of our theory by demonstrating how disastrous results are possible when the LGR/PS and LG/PS methods are artificially forced to solve BVP-type problems over a finite horizon. Furthermore, in practical optimal control problems, this disaster may be masked for low orders of discretization by giving the appearance of convergence;however, as the mesh is refined, the LGR and LG PS methods reveal a lack of convergence, particularly at the boundary points similar to the classic Runge phenomenon.[42] Our theory explains this phenomenon thus suggesting that the weighted-interpolant perspective is indeed the proper perspective for all PS methods for optimal control.

II.   A Distilled Optimal Control Problem

For simplicity in exposition, we will consider the following scalar Bolza problem, with the understanding that our discretization methods can be easily extended to higher-dimensional cases. For the same reason we will ignore path constraints as these can also be easily incorporated into our framework by replacing the control Hamiltonian by the Lagrangian of the Hamiltonian.[26] In simplifying such bookkeeping issues, we consider the problem of finding the optimal state-control function pair $t \mapsto (x, u) \in \mathbb{R} \times \mathbb{R}$ that solves the following problem (Problem B):

$$x \in \mathbb{X} \subseteq \mathbb{R}, \quad u \in \mathbb{R}$$

$$(B) \begin{cases} \text{Minimize} & J[x(\cdot), u(\cdot)] = E(x(-1), x(1)) + \int_{-1}^{1} F(x(t), u(t)) \, dt \\ \text{Subject to} & \dot{x}(t) = f(x(t), u(t)) \\ & e(x(-1), x(1)) = 0 \end{cases}$$

where $\mathbb{X}$ is an open set in $\mathbb{R}$ and the problem data (i.e. the endpoint cost function, E, the running cost function, F, the vector field, f and the endpoint constraint function, e) are assumed to be at least C^1-smooth.

An application of Pontryagin's Minimum Principle (PMP) results in a two-point BVP that we denote as Problem B^λ. This "dualization" is achieved through the construction of the control Hamiltonian, H, and endpoint Lagrangian, $\bar{E}$, defined as

$$H(\lambda, x, u) = F(x, u) + \lambda f(x, u) \tag{1}$$
$$\bar{E}(\nu, x(-1), x(1)) = E(x(-1), x(1)) + \nu e(x(-1), x(1)) \tag{2}$$

where λ is the adjoint covector (or costate) and ν is the endpoint covector. Problem B^λ is now posed as follows:

$$(B^\lambda) \begin{cases} \text{Find} & t \mapsto (x, u, \lambda) \text{ and } \nu \\ \text{Such that} & \dot{x}(t) - f(x(t), u(t)) = 0 \\ & e(x(-1), x(1)) = 0 \\ & \dot{\lambda}(t) + \partial_x H(\lambda(t), x(t), u(t)) = 0 \\ & \partial_u H(\lambda(t), x(t), u(t)) = 0 \\ & \lambda(-1) + \dfrac{\partial \bar{E}}{\partial x(-1)}(\nu, x(-1), x(1)) = 0 \\ & \lambda(1) - \dfrac{\partial \bar{E}}{\partial x(1)}(\nu, x(-1), x(1)) = 0 \end{cases}$$

In using the PMP to solve a trajectory optimization problem, we must solve Problem B^λ which is clearly a problem of finding the zeros of a map in an appropriate function space.[43] This is the so-called indirect method. From a first-principles perspective,[31] the PMP is used as follows: For every optimal solution to Problem B, there exist an adjoint covector function, $t \mapsto \lambda$, and an endpoint covector, ν, that satisfy the conditions set forth by Problem B^λ. Thus, every candidate optimal solution to Problem B must also be a solution to Problem B^λ under appropriate technical conditions. The LGL/PS method ensures a satisfaction of this condition through its Covector Mapping Theorem.[25, 26] A generalized and unified version of this theorem with respect to PS methods is the main contribution of this paper and is developed in the following sections. A generalized version of this theorem with respect to the stability and convergence of the approximation is described by Gong et al.[28]

III.   A Unified Perspective on Pseudospectral Methods

Legendre PS methods are a subset of a larger class of spectral methods which use orthogonal basis functions in global expansions similar to Fourier and Sinc series expansions.[44] What distinguishes PS methods is that they are based on approximating the unknown functions by interpolants.[45] This is one reason why PS methods are sharply different from other polynomial methods previously considered in the literature. The interpolating points in a PS method can be any of the three major classes of Gaussian quadrature points called nodes: Gauss points, Radau points and Lobatto points. Note that Radau and Lobatto points are also Gaussian quadrature points; this is why specific PS methods are identified by adjectives that describe the specific choice of nodes. All of these nodes are zeros of the orthogonal polynomials or their derivatives such as Legendre and Chebyshev or more generally the Jacobi polynomials. Thus, a very large family of PS methods can be generated to solve optimal control problems. This notion is similar to Runge-Kutta (RK) methods. For example, Euler, trapezoid, Hermite-Simpson and many other methods are equivalently some form of an RK method with appropriately chosen coefficients.[46] Just as not all RK methods are legitimate, not all PS methods are legitimate. This is particularly true in solving optimal control problems as has been noted by Hager.[43] For example, an RK method that is legitimate for propagating an ODE may fail gloriously when applied to an optimal control problem. In other words, it is imperative to carefully select an appropriate method to solve optimal control problems.

In the absence of special information (such as special boundary conditions), it is customary to select Chebyshev PS methods in solving PDE problems as they are simple and effective. For optimal control of ODEs, the LGL/PS method is preferable as it maintains the consistency of dual approximations similar to the Hager class of RK methods. Recently, we proposed the usage of Radau points[38, 39, 41] as a means to manage a singularity problem that arises in solving infinite-horizon control problems. The goal of this paper is to present a more unified view of these methods where we can show the similarities and differences of the

methods based on the choice of these different nodes and show the effect of these nodes on the discretization of both the primal and dual problems.

A.   Weighted Interpolants and Differentiation Matrices on Arbitrary Grids

We begin with the basic definition of PS methods and the different ingredients required in defining these methods for various nodes. Inspired by Weideman's research,[45] we select PS methods based on weighted interpolants of the form,

$$y(t) \approx y^N(t) = \sum_{j=0}^{N} \frac{W(t)}{W(t_j)} \phi_j(t) y_j, \ a \le t \le b \tag{3}$$

where $y(t)$ is an arbitrary function. Here the nodes $t_j, j = 0, ..., N$ are a set of distinct interpolation nodes (defined below) in the interval $[a, b]$, the weight function $W(t)$ is a positive function on the interval, and $\phi_j(t)$ is the $Nth-$ order Lagrange interpolating polynomial that satisfies the relationship $\phi_j(t_k) = \delta_{jk}$. This implies that

$$y_j = y^N(t_j), \ \ j = 0, ...N. \tag{4}$$

An expression for the Lagrange polynomial can be written as[44]

$$\phi_j(t) = \frac{g_N(t)}{g_N'(t_j)(t - t_j)}, \ \ g_N(t) = \prod_{j=0}^{N}(t - t_j). \tag{5}$$

One important tenant of polynomial approximation of functions is that differentiation of the approximated functions can be performed by differentiation of the interpolating polynomial,

$$\frac{dy^N(t)}{dt} = \sum_{j=0}^{N} \frac{y_j}{W(t_j)}[W'(t)\phi_j(t) + W(t)\phi_j']$$

Since only the values of the derivative at the nodes t_i are required for PS methods, then we have,

$$\frac{dy^N(t)}{dt}\bigg|_{t_i} = \sum_{j=0}^{N} \frac{y_j}{W(t_j)}[W'(t_i)\delta_{ij} + W(t_i)D_{ij}] = \sum_{j=0}^{N} D_{ij}[W]y_j \tag{6}$$

where we use $D_{ij}[W]$ as a shorthand notation for the W-weighted differentiation matrix,

$$D_{ij}[W] := \frac{[W'(t_i)\delta_{ij} + W(t_i)D_{ij}]}{W(t_j)} \tag{7}$$

and D_{ij} is usual unweighted differentiation matrix given by,

$$D_{ij} := \frac{d\phi_j(t)}{dt}\bigg|_{t=t_i} \tag{8}$$

Thus, when $W(t) = 1$, we have

$$D_{ij}[1] = D_{ij} \tag{9}$$

From Eq. (5), the unweighted differentiation matrix, $D_{ij} = \phi_j'(t_i)$, has the form,

$$D_{ij} = \begin{cases} \dfrac{g_N'(t_i)}{g_N'(t_j)} \dfrac{1}{(t_i - t_j)}, & i \ne j \\[4mm] \dfrac{g_N''(t_i)}{2g_N'(t_i)}, & i = j \end{cases} \tag{10}$$

The above equations are the general representations of the derivative of the Lagrange polynomials evaluated at arbitrary interpolation nodes. Thanks to Runge, it is well-known that an improper selection of the grid points can lead to disastrous consequences. In fact, a uniform distribution of grid points is the worst possible choice for polynomial interpolation and hence differentiation. On the other hand, the best possible choice of grid points for integration, differentiation and interpolation of functions are Gaussian quadrature points.[36] Consequently, all PS methods use Gaussian quadrature points. In order to limit the scope of this paper, we limit our discussions to Gaussian points based on Legendre polynomials as these polynomials offer elegant theoretical properties for optimal control applications.

American Institute of Aeronautics and Astronautics

B. Legendre Nodes

Gaussian quadrature points lie in the interval $[-1, 1]$. In most problems the physical problem is posed on some interval $[a, b]$ which can then be easily transformed to the computational domain $[-1, 1]$ via a linear transformation. The Gauss quadrature points are zeros that are interior to the interval $[-1, 1]$. The Gauss-Radau zeros include one of the end points of the interval, usually the left-end point at $t = -1$. The Gauss-Lobatto points include both endpoints of the interval at $t = -1$, and $t = 1$; see Fig. 1. These
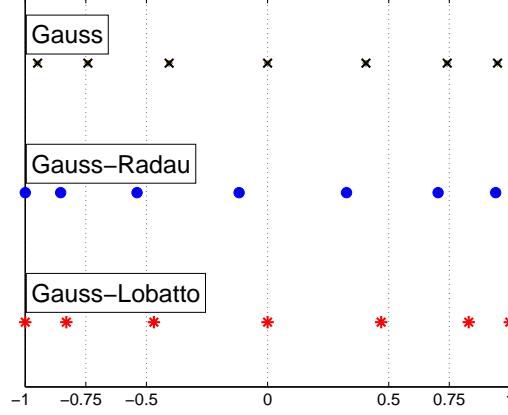

Figure 1. Illustrating the various Legendre quadrature nodes: Legendre-Gauss (LG), Legendre-Gauss-Radau (LGR) and Legendre-Gauss-Lobatto (LGL) grid points.

quadrature nodes are related to the zeros of the Qth-order Jacobi polynomials $P_Q^{\alpha,\beta}$ which are orthogonal on the interval $[-1, 1]$ with respect to the inner product,[36, 37]

$$\int_{-1}^{1} (1-t)^{\alpha}(1+t)^{\beta} P_Q^{\alpha,\beta}(t) P_{Q'}^{\alpha,\beta}(t) dt \tag{11}$$

Note that the Legendre polynomials are a special case of Jacobi polynomials which correspond to $\alpha = \beta = 0$. Let $t_{i,Q}^{\alpha,\beta}$ be the Q zeros of the Qth-order Jacobi polynomial $P_Q^{\alpha,\beta}$ such that

$$P_Q^{\alpha,\beta}(t_{i,Q}^{\alpha,\beta}) = 0, \quad i = 0, ..., Q-1 \tag{12}$$

Then, we can define zeros and weights that approximate the following integral

$$\int_{-1}^{1} u(t) dt = \sum_{i=0}^{N} w_i u(t_i) + R(u) \tag{13}$$

and find expressions for the unweighted differentiation matrices for these nodes as follows:[37]

1. Legendre-Gauss (LG)

$$t_i = t_{i,N+1}^{0,0}, \quad i = 0, \dots, N$$

$$w_i^{0,0} = \frac{2}{1-(t_i)^2} \left[\frac{d}{dt} (L_{N+1}(t)|_{t=t_i}) \right]^{-2}, \quad i = 0, \dots, N,$$

$$R(u) = 0 \ \text{ if } \ u(t) \in \mathcal{P}^{2N+1}([-1, 1]).$$

$$D_{ij} = \begin{cases} \frac{L'_{N+1}(t_i)}{L'_{N+1}(t_j)(t_i - t_j)}, & i \neq j, 0 \leq i, j \leq N \\ \frac{t_i}{1-t_i^2}, & i = j \end{cases}$$

American Institute of Aeronautics and Astronautics

2. *Legendre-Gauss-Radau (LGR)*

$$t_i = \begin{cases} -1, & i = 0, \\ t^{0,1}_{i-1,N}, & i = 0, \ldots, N \end{cases}$$

$$w_i^{0,0} = \frac{1 - t_i}{(N+1)^2 [L_N(t_i)]^2}, \ i = 0, \ldots, N,$$

$$R(u) = 0 \ \text{ if } \ u(t) \in \mathcal{P}^{2N}([-1,1]).$$

$$D_{ij} = \begin{cases} \frac{-N(N+2)}{4}, & i = j = 0, \\ \frac{L_N(t_i)}{L_N(t_j)} \frac{1-t_j}{1-t_i} \frac{1}{(t_i-t_j)}, & i \neq j, 0 \leq i,j \leq N \\ \frac{1}{2(1-t_i)}, & 1 \leq i = j \leq N \end{cases}$$

3. *Legendre-Gauss-Lobatto (LGL)*

$$t_i = \begin{cases} -1, & i = 0, \\ t^{1,1}_{i-1,N-1}, & i = 1, \ldots, N-1, \\ 1 & i = 1 \end{cases}$$

$$w_i^{0,0} = \frac{2}{N(N+1)[L_N(t_i)]^2}, \ i = 0, \ldots, N,$$

$$R(u) = 0 \ \text{ if } \ u(t) \in \mathcal{P}^{2N-1}([-1,1]).$$

$$D_{ij} = \begin{cases} \frac{-N(N+1)}{4}, & i = j = 0, \\ \frac{L_N(t_i)}{L_N(t_j)} \frac{1}{(t_i-t_j)}, & i \neq j, 0 \leq i,j \leq N \\ 0, & 1 \leq i = j \leq N-1 \\ \frac{N(N+1)}{4}, & i = j = N, \end{cases}$$

where $\mathcal{P}^N$ is the set of all algebraic polynomials of degree at most N.

C. Pre-Hilbert Spaces

Given the definitions of the quadrature nodes and weights above, we can proceed with the definitions of the appropriate discrete inner-product spaces associated with the above nodes and weights. Let $[-1,1] \mapsto \{y(t), z(t)\}$ be real-valued functions in $L^2([-1,1], \mathbb{R})$. The *standard* inner product in L^2 is given by

$$\langle y, z \rangle_{L^2} := \int_{-1}^{+1} y(t)z(t) \, dt$$

For any given $m \in \mathbb{N}$, and $\{w_j > 0, \ j = 0, 1, \ldots m\}$, define the *weighted discrete* inner product in L^2 to be

$$\langle y, z \rangle_{m,w} := \sum_{j=0}^{m} y(t_j) w_j z(t_j)$$

For the various Gauss quadrature nodes and the standard inner-product in L^2, we have the following result:[44]

Lemma 1 *For all $pq \in \mathcal{P}^{2N+\zeta}$,*

$$\langle p, q \rangle_{L^2} = \langle p, q \rangle_{N,w}$$

where $\zeta = 1$ for Legendre-Gauss (LG), $\zeta = 0$ Legendre-Gauss-Radau (LGR) and $\zeta = -1$ for Legendre-Gauss-Lobatto (LGL) integration and weights.

American Institute of Aeronautics and Astronautics

For any weight function, $W(t)$, the *weighted* inner product in L^2 is given by

$$\langle y, z \rangle_{L_W^2} := \int_{-1}^{+1} y(t) W(t) z(t) \, dt$$

Let $W(t)$ be any one of the following weight functions,

$$W(t) = \begin{cases} W_{lgl}(t) & \Rightarrow W(t) = 1 \\ W_{lgr}(t) & \Rightarrow W(t) = 1 - t \\ W_{lg}(t) & \Rightarrow W(t) = 1 - t^2 \end{cases} \tag{14}$$

From Lemma 1, we have the following unified result:

Lemma 2 *For all $pq \in \mathcal{P}^{2N-1}$,*

$$\langle p, q \rangle_{L_W^2} = \langle p, q \rangle_{N,w}$$

Remark 1 *Lemma 2 implies that W_{lgl} is the only weight function that is dual to itself in the standard inner product space and unweighted interpolation.*

This primal-dual symmetry explains why the LGL PS method offers elegant properties.

Lemma 3 *For the three Legendre quadrature nodes, LG, LGR and LGL, the following relationships hold: For $i \neq j$ we have,*

$$D_{ij} = -\frac{W(t_j)}{W(t_i)} \frac{w_j}{w_i} D_{ji}$$

Hence, we can write,

$$D_{ij}[W] = -\frac{w_j}{w_i} D_{ji}, \ i \neq j$$

and for $i = j$ the following relationships hold:

$$\text{For LG Nodes} \quad \rightarrow \quad D_{ii}[W_{lg}] = -D_{ii}, i = 0, \cdots, N \tag{15}$$

$$\text{For LGR Nodes} \quad \rightarrow \quad \begin{cases} D_{ii}[W_{lgr}] = -D_{ii} \ i = 1, \cdots, N \\ D_{00}[W_{lgr}] = -D_{00} - \frac{1}{w_0} \end{cases} \tag{16}$$

$$\text{For LGL Nodes} \quad \rightarrow \quad \begin{cases} D_{ii}[W_{lgl}] = D_{ii} = 0, \ \ i = 1, \cdots, N-1 \\ D_{00}[W_{lgl}] = -D_{00} - 1/w_0 \\ D_{NN}[W_{lgl}] = -D_{NN} + 1/w_N \end{cases} \tag{17}$$

IV. Primal-Dual Pseudospectral Discretizations

We can now define LG, LGR and LGL pseudospectral methods based on the choice of interpolation nodes. For approximating the state and costate variables we use the expansions,

$$x^N(t) = \sum_{j=0}^{N} \frac{W(t)}{W(t_j)} x_j \phi_j(t), \quad \lambda^N(t) = \sum_{j=0}^{N} \frac{W^*(t)}{W^*(t_j)} \lambda_j \phi_j(t) \tag{18}$$

where W and W^* are appropriate choices of weight functions. Lemma 2 suggests that if we take $W(t) = 1$ and LGL nodes, then we must take $W^*(t) = 1$ as well. This is in conformance with the standard theory.[26] It is clear that the use of Lemma 2 in the integration-by-parts argument proposed in Refs. [26, 56], generalizes the standard theory by suggesting that if we take $W(t) = 1$ and LGR nodes, then $W^*(t) = 1 - t$; likewise, for $W(t) = 1$ and LG nodes, $W^*(t) = 1 - t^2$. Similarly, if we take $W(t) = 1 - t$ for LGR nodes, then $W^*(t) = 1$ etc.; thus,

$$(W^*(t))^* = W(t) \tag{19}$$

That is, the dual of the dual is the original function.

American Institute of Aeronautics and Astronautics

For approximating the control variables, we take

$$u^N(t) = \sum_{j=0}^{N} u_j \psi_j(t) \tag{20}$$

where $\psi_j(t)$ is any interpolant. *Note this key point that constitutes part of the unification principles*[28, 30]. That is $u^N(t)$ is not necessarily constructed as a Lagrange interpolating polynomial; rather, $u^N(t)$ is constructed using any type of interpolation such as linear, cubic, spline etc. In previous versions, it was assumed that $u^N(t)$ must be Lagrange interpolants leading to certain convergence questions with increase in N. By choosing $u^N(t)$ to be any interpolant, we retain the key elements of pseudospectral theory while allowing $\psi_j(t)$ to be chosen arbitrarily. This arbitrariness can then be exploited to further enhance the spectral convergence property of PS methods by collecting information in between nodes through an application of Bellman's Principle. Thus ψ_j becomes an abstract interpolant that may be selected based on optimization principles. These ideas have been fully discussed by Ross et al.[47]

The preceding points illustrate how PS methods for optimal control have indeed evolved to a separate and distinct theory from those used to solve problems in fluid dynamics, the original source of ideas for PS methods for optimal control. It is clear that this paper is part of these new foundations for PS methods for optimal control.

A.   Discretization of the Primal Problem (B^N)

Based on these new ideas, we formulate PS methods for optimal control as follows: From Eqs. (6) and (18), it follows that the state derivative is approximated as,

$$\left. \frac{dx^N(t)}{dt} \right|_{t_i} = \sum_{j=0}^{N} D_{ij}[W]x_j \tag{21}$$

Suppose we set $W(t) \equiv 1$ for the state variables for all Legendre PS methods. This automatically defines dual weight functions for costate interpolation. We may also set $W(t)$ according to the appropriate class of weighted interpolations in which case we would arrive at the corresponding duals. Choosing $W(t) \equiv 1$ for the primal variables implies that the primal differentiation matrix, $D[W]$, is the standard differentiation matrix (Cf. Eq. (9)). Applying the quadrature rules for approximating the integral terms and using the appropriate form of the derivative matrix according to the choice of the nodes, we have the following expressions for Problem B^N which is the discretized form of Problem B:

$$X \in \mathbb{R}^{N_n}, \quad U \in \mathbb{R}^{N_n}$$

$$(B^N) \begin{cases} \text{Minimize} & J^N[X,U] = E(x_0, x_N) + \sum_{j=0}^{N} w_j F(x_j, u_j) \\ \\ \text{Subject to} & \sum_{j=0}^{N} D_{ij}x_j - f(x_i, u_i) = 0 \quad i = 0, 1, \ldots, N \\ \\ & e(x_0, x_N) = 0 \end{cases}$$

where $X = [x_0, x_1, \ldots, x_{N_n}]$ and $U = [u_0, u_1, \ldots, u_{N_n}]$.

B.   Selection of a Proper Choice of PS Methods

Lemma 2 suggests the proper choice of an appropriate PS method for Problem B^N. In many optimal control problems the endpoints are not totally free and are frequently constrained; that is, the endpoint function, e, is typically a function of both $x(-1)$ and $x(1)$. Then, according to Lemma 2, we must choose the LGL/PS method. Now suppose that we choose an LG or LGR PS method for such a problem. The motivation for this selection may be based on Lemma 1 which suggests the apparently higher degree of approximation for LG and LGR nodes when compared to LGL nodes. That this is not true is best illustrated by a simple

counter example of Gong et al:[30]

$$\mathbf{x} = (x, v) \in \mathbb{R}^2, \quad u \in \mathbb{R}$$

$$(P_1) \begin{cases} \text{Minimize} & J[\mathbf{x}(\cdot), u(\cdot)] = \displaystyle\int_0^1 v(t)u(t) \, dt \\ \text{Subject to} & \dot{x}(t) = v(t) \\ & \dot{v}(t) = -v(t) + u(t) \\ & v(t) \geq 0 \\ & 0 \leq u(t) \leq 2 \\ & (x(0), v(0)) = (0, 1) \\ & (x(1), v(1)) = (1, 1) \end{cases}$$

This problem describes a particle of unit mass $(m = 1)$ moving in a linear resistive medium $(\alpha = -1)$ where
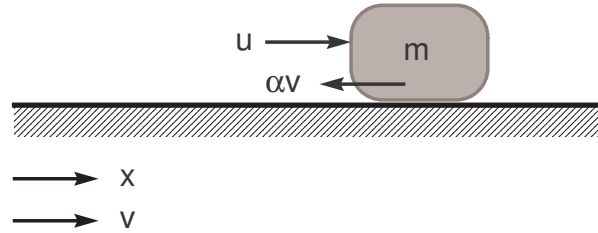

Figure 2. Schematic for Problem P_1.

x is the position, v is the velocity and u is the applied force; see Fig. 2. The optimal control problem is to minimize the total amount of work done. From physical considerations or by a direct application of the Minimum Principle, it is easy to verify that the optimal control is a constant and is equal to the amount of force required to maintain the initial speed. Thus, the exact optimal control is given by,

$$u^*(t) = 1 \tag{22}$$

Since the optimal control is a polynomial, a correct PS method is expected to achieve exact performance for sufficiently large N. The solution to this problem based on the standard LGL discretization is shown in Fig. 3 for 10 nodes. Predictably, the LGL PS method generates the exact solution. On the other hand, the
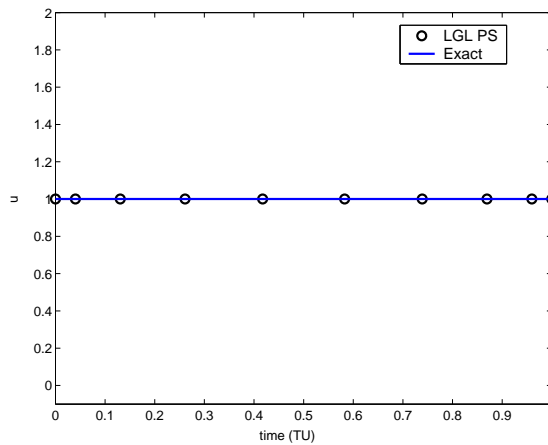

Figure 3. Control solution to Problem P_1 for $N = 10$: LGL PS method.

LG and LGR PS methods produce disastrous solutions as shown in Fig. 4. Thus, choosing a PS method for
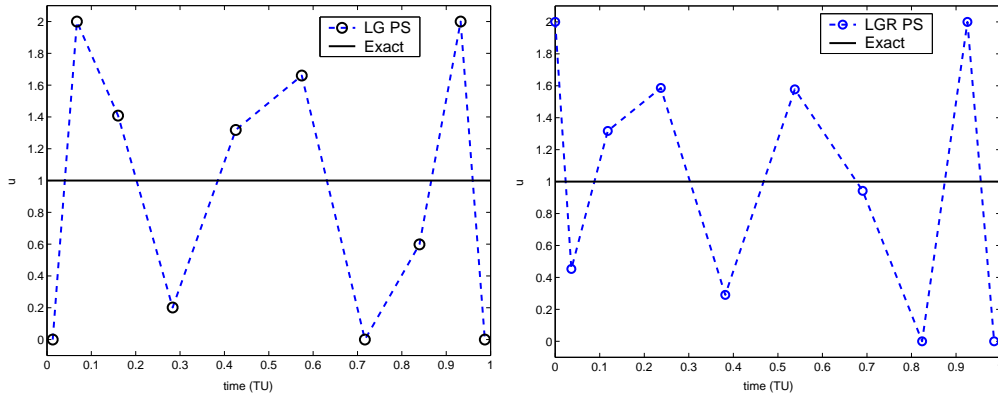
American Institute of Aeronautics and Astronautics

Figure 4. Control solutions to Problem P_1 for $N = 10$: LGR and LG PS methods; compare Fig. 3.

optimal control based on Lemma 1 alone is a bad proposition. To illustrate the point that the LG and LGR PS methods are not just converging at a slower rate than the LGL PS method, their control solutions are plotted in Fig. 5 for $N = 30$. It is clear that the LG and LGR PS methods do not converge. What is most
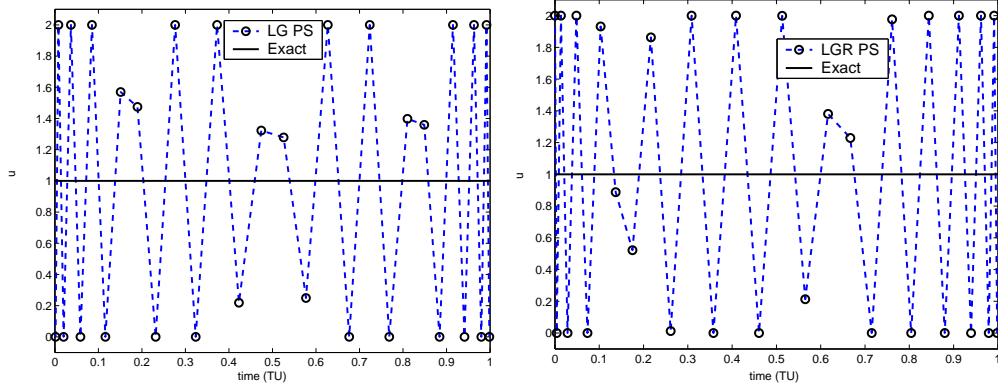

Figure 5. Control solutions to Problem P_1 for $N = 30$: Illustrating why the LGR and LG PS methods are the wrong PS methods for this problem.

interesting about the lack of convergence of the LGR and LG PS methods is that their node points converge to the node points of the LGL PS method as N increases. This point is illustrated in Fig. 6.

C. Discretization of the Dualized Problem ($B^{\lambda N}$)

Given that dual space considerations are indeed important even when only primal variables are sought,[34, 43] we now explore the discretization of the dualized problem. From Eqs. (7) and (18), it follows that the costate derivative is approximated as,

$$\left.\frac{d\lambda^N(t)}{dt}\right|_{t_i} = \sum_{j=0}^{N} D_{ij}^* \lambda_j \tag{23}$$

with

$$D_{ij}^* = D_{ij}[W] = \frac{[W'(t_i)\delta_{ij} + W(t_i)D_{ij}]}{W(t_j)}$$
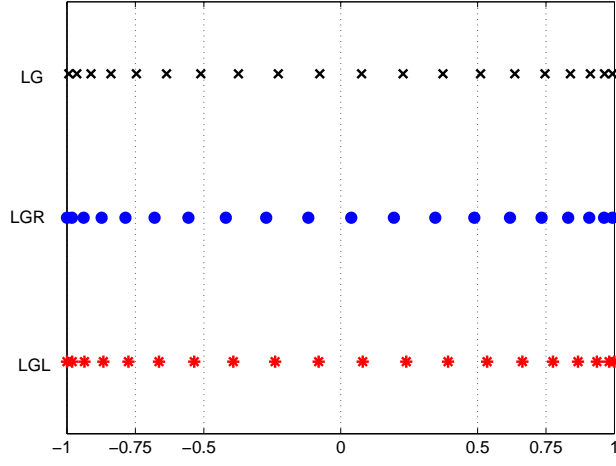
Figure 6. Illustrating the near absence of differences between the three Legendre quadrature nodes for increasing N.

Thus Problem $B^{\lambda N}$ which is the discretization of Problem B^{λ} defined in Sec. II, can be constructed as,

$$
(B^{\lambda N})\begin{cases}
\displaystyle\sum_{j=0}^{N} D_{ij}x_j - f(x_i, u_i) &= 0 \qquad i = 0, 1, \ldots, N \\[2ex]
e(x_0, x_N) &= 0 \\[2ex]
\displaystyle\sum_{j=0}^{N} D^*_{ij}\lambda_j + \partial_{x_i} H(\lambda_i, x_i, u_i) &= 0 \qquad i = 0, 1, \ldots, N \\[2ex]
\partial_{u_i} H(\lambda_i, x_i, u_i) &= 0 \qquad i = 0, 1, \ldots, N \\[2ex]
\lambda_0 + \dfrac{\partial \bar{E}}{\partial x_0}(\nu, x_0, x_N) &= 0 \\[2ex]
\lambda_N - \dfrac{\partial \bar{E}}{\partial x_N}(\nu, x_0, x_N) &= 0
\end{cases}
$$

Thus, in the unified PS method, we use the matrices D and D^* as primal and dual differentiation matrices arising from the primal and dual weighted interpolation with weight functions, $W(t)$ and $W^*(t)$, respectively. Furthermore, from Eq. (19), it follows that we may choose $W(t)$ for discretizing the dual, in which case we need to use $W^*(t)$ for discretizing the primals.

V. A Unified Covector Mapping Theorem

Based on the preceding discussions it is now abundantly clear that the correct inner product space for all three PS discretizations is given by $\mathbb{R}^{N_n}_w$, the finite-dimensional Hilbert space $\mathbb{R}^{N_n}$ equipped with the *weighted* inner product,

$$
\langle a, b\rangle_{\mathbb{R}^{N_n}_w} := \sum_{i=0}^{N} a_i w_i b_i \qquad a, b \in \mathbb{R}^{N_n}
$$

Thus, the Lagrangian in $\mathbb{R}^{N_n}_w$ is given by,[40]

$$
\begin{aligned}
\overline{J^N}[\widetilde{\lambda}, \widetilde{\nu}, x, u] &= \sum_{j=0}^{N} w_j F(x_j, u_j) + \sum_{i=0}^{N} w_i \widetilde{\lambda}_i f(x_i, u_i) - \sum_{i=0}^{N} w_i \widetilde{\lambda}_i \sum_{j=0}^{N} D_{ij}x_j + \widetilde{\nu}e(x_0, x_N) + E(x_0, x_N) \\[2ex]
&= \sum_{i=0}^{N} w_i H(\widetilde{\lambda}_i, x_i, u_i) - \sum_{i=0}^{N} w_i \widetilde{\lambda}_i \sum_{j=0}^{N} D_{ij}x_j + \bar{E}(\widetilde{\nu}, x_0, x_N)
\end{aligned}
$$

American Institute of Aeronautics and Astronautics

where $\widetilde{\lambda}$ and $\widetilde{\nu}$ are the Karush-Kuhn-Tucker (KKT) multipliers in $\mathbb{R}_w^{N_n}$. From the KKT theorem we have,

$$
\begin{aligned}
\partial_{x_k} \overline{J^N}[\widetilde{\lambda}, \widetilde{\nu}, x, u] &= w_k \partial_{x_k} H(\widetilde{\lambda}_k, x_k, u_k) - \sum_{i=0}^N w_i \widetilde{\lambda}_i \sum_{j=0}^N D_{ij} \delta_{jk} \qquad k = 1, \ldots, N-1 \\
&= w_k \partial_{x_k} H(\widetilde{\lambda}_k, x_k, u_k) - \sum_{i=0}^N w_i \widetilde{\lambda}_i D_{ik}
\end{aligned}
$$

From Lemma 3 we have,

$$
D^*_{ki} := -\frac{w_i}{w_k} D_{ik} \quad for\ k = 1, \ldots, N-1, i = 0, \ldots, N \tag{24}
$$

Hence,

$$
\partial_{x_k} \overline{J^N}[\widetilde{\lambda}, \widetilde{\nu}, X, U] = w_k \left(\partial_{x_k} H(\widetilde{\lambda}_k, x_k, u_k) + \sum_{i=0}^N D^*_{ki} \widetilde{\lambda}_i \right) = 0
$$

Similarly,

$$
\begin{aligned}
\partial_{x_k} \overline{J^N}[\widetilde{\lambda}, \widetilde{\nu}, X, U] &= w_k \partial_{x_k} H(\widetilde{\lambda}_k, x_k, u_k) - \sum_{i=0}^N w_i \widetilde{\lambda}_i \sum_{j=0}^N D_{ij} \delta_{jk} + \widetilde{\nu} \partial_{x_k} e(x_0, x_N) \qquad k = 0, N \\
&= w_k \partial_{x_k} H(\widetilde{\lambda}_k, x_k, u_k) - \sum_{i=0}^N w_i \widetilde{\lambda}_i D_{ik} + \widetilde{\nu} \partial_{x_k} e(x_0, x_N)
\end{aligned}
$$

From Lemma 3, this implies,

$$
\partial_{x_k} \overline{J^N}[\widetilde{\lambda}, \widetilde{\nu}, X, U] = w_k \left(\partial_{x_k} H(\widetilde{\lambda}_k, x_k, u_k) + \sum_{i=0}^N D^*_{ki} \widetilde{\lambda}_i \right) + \widetilde{\lambda}^*_k + \widetilde{\nu} \partial_{x_k} e(x_0, x_N) = 0, \quad k = 0
$$

$$
\partial_{x_k} \overline{J^N}[\widetilde{\lambda}, \widetilde{\nu}, X, U] = w_k \left(\partial_{x_k} H(\widetilde{\lambda}_k, x_k, u_k) + \sum_{i=0}^N D^*_{ki} \widetilde{\lambda}_i \right) + \widetilde{\lambda}^*_k - \widetilde{\nu} \partial_{x_k} e(x_0, x_N) = 0, \quad k = N
$$

where $\widetilde{\lambda}^*_k$ is given by:

1. *Legendre-Gauss*

$\widetilde{\lambda}^*_k = 0$ for $k = 0$ and N

2. *Legendre-Gauss-Radau*

$\widetilde{\lambda}^*_k = \widetilde{\lambda}_k$ for $k = 0$ and $\widetilde{\lambda}^*_k = 0$ for $k = N$

3. *Legendre-Gauss-Lobatto*

$\widetilde{\lambda}^*_k = \widetilde{\lambda}_k$ for $k = 0$ and N

Thus Problem $B^{N\lambda}$ can be written as,

$$
(B^{N\lambda}) \begin{cases}
\displaystyle\sum_{j=0}^{N} D_{ij} x_j - f(x_i, u_i) & = 0 \qquad i = 0, 1, \ldots, N \\[2ex]
e(x_0, x_N) & = 0 \\[2ex]
\displaystyle\sum_{j=0}^{N} D_{ij}^* \widetilde{\lambda}_j + \partial_{x_i} H(\widetilde{\lambda}_i, x_i, u_i) & = 0 \qquad i = 1, \ldots, N-1 \\[2ex]
\partial_{u_i} H(\widetilde{\lambda}_i, x_i, u_i) & = 0 \qquad i = 0, 1, \ldots, N \\[2ex]
\displaystyle\sum_{j=0}^{N} D_{ij}^* \widetilde{\lambda}_j + \partial_{x_i} H(\widetilde{\lambda}_i, x_i, u_i) & = -c_0^* \qquad i = 0 \\[2ex]
\displaystyle\sum_{j=0}^{N} D_{ij}^* \widetilde{\lambda}_j + \partial_{x_i} H(\widetilde{\lambda}_i, x_i, u_i) & = -c_N^* \qquad i = N \\[2ex]
\widetilde{\lambda}_0^* + \dfrac{\partial \bar{E}}{\partial x_0}(\widetilde{\nu}, x_0, x_N) & = w_0 c_0^* \\[2ex]
\widetilde{\lambda}_N^* - \dfrac{\partial \bar{E}}{\partial x_N}(\widetilde{\nu}, x_0, x_N) & = w_N c_N^*
\end{cases}
$$

where c_0^* and c_N^* are arbitrary real numbers.

Now suppose that a solution to Problem $B^{\lambda N}$ exists. That is, we suppose that a (discretized) solution to the two-point BVP exists. Then the covectors of Problem $B^{\lambda N}$ provide an existence theorem for the solution of Problem $B^{N\lambda}$; that is, for all $k = 0, \ldots, N$, $\widetilde{\lambda}_k = \lambda_k$ and $\widetilde{\nu} = \nu$ under the following conditions, a solution to Problem $B^{N\lambda}$ exists:

Condition A:

Legendre-Gauss

x_0 and x_N are free; i.e. unspecified.

Legendre-Gauss-Radau

x_0 may be arbitrarily specified but x_N must be free, or vice versa.

Legendre-Gauss-Lobatto

Both x_0 and x_N may be arbitrarily specified.

The dual to Condition A is given by,

Condition A^*:

Legendre-Gauss

$x_0 = 0 = x_N$.

Legendre-Gauss-Radau

x_0 may be arbitrarily specified but $x_N = 0$, or vice versa.

Legendre-Gauss-Lobatto

Both x_0 and x_N may be arbitrarily specified.

Thus, any solution to Problem $B^{\lambda N}$ is also a solution to Problem $B^{N\lambda}$; in this case, we have,

$$
c_0^* = 0 = c_N^* \tag{25}
$$

That is, Eq. (25) is a matching condition that matches Problem $B^{N\lambda}$ to Problem $B^{\lambda N}$. Matching conditions[48, 49] are part of the totality of closure conditions required to complete the circuit (arrows) indicated in Fig. 7.

Note that Problems B^N and $B^{\lambda N}$ are generated from Problems B and B^λ respectively without introducing any additional continuous-time primal conditions to carry over to the discrete-time problems. This notion is implicit in Fig. 7. Fig. 7 is a commutative diagram promulgated in Ref. [27] that forms part of the broader
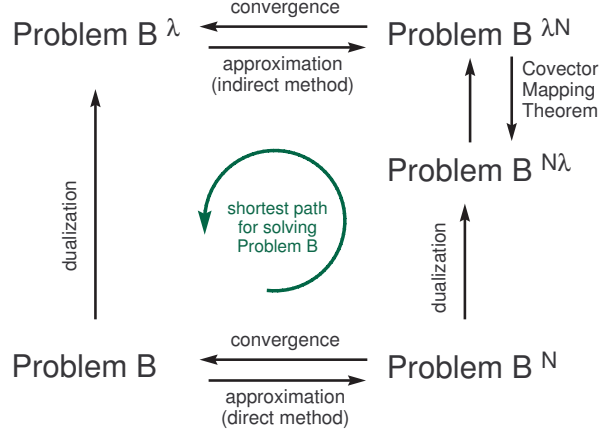

Figure 7. Illustrating the Covector Mapping Principle and the unification of direct and indirect methods in optimal control as first described by Ross and Fahroo.[25-27]

notion of the Covector Mapping Principle in optimal control. A historical account of the origins of this diagram and the ideas embedded in it are documented in Ref. [33]. *The concepts embedded in Fig. 7 form yet another unification principle in PS methods and optimal control at large.*

A. A Unified Pseudospectral Covector Mapping Theorem

In completing the steps suggested in Fig. 7 towards the development of a Unified Covector Mapping Theorem, we identify the following multiplier sets analogous to those introduced in Ref. [50]: Let $\chi := \{[\mathbf{x}_k], [\mathbf{u}_k]\}$ and $\Lambda := \{\boldsymbol{\nu}_0, [\boldsymbol{\mu}_k], [\boldsymbol{\lambda}_k]\}$. We denote by $\mathbb{M}^{\lambda N}(\chi)$ the multiplier set corresponding to χ,

$$\mathbb{M}^{\lambda N}(\chi) := \left\{ \Lambda : \Lambda \text{ satisfies conditions of Problem } B^{\lambda N} \right\} \tag{26}$$

Similarly, we define, $\widetilde{\Lambda} := \left\{ \widetilde{\nu}_0, [\widetilde{\boldsymbol{\mu}}_k], [\widetilde{\boldsymbol{\lambda}}_k] \right\}$ and $M^{N\lambda}(\chi)$ the multiplier set,

$$\mathbb{M}^{N\lambda}(\chi) := \left\{ \widetilde{\Lambda} : \widetilde{\Lambda} \text{ satisfies conditions of Problem } B^{N\lambda} \right\} \tag{27}$$

Clearly, $\mathbb{M}^{\lambda N}(\chi) \subseteq \mathbb{M}^{N\lambda}(\chi)$. We now define a new multiplier set,

$$\widehat{\mathbb{M}}^{N\lambda}(\chi) := \left\{ \widetilde{\Lambda} \in \mathbb{M}^{N\lambda}(\chi) : \widetilde{\Lambda} \text{ satisfies Eq. (25)} \right\} \tag{28}$$

Thus, $\widehat{\mathbb{M}}^{N\lambda}(\chi) \sim \mathbb{M}^{\lambda N}(\chi)$. That is, under a matching (closure) condition, every solution of Problem $B^{N\lambda}$ is also a solution to Problem $B^{\lambda N}$. This statement is encapsulated as the Covector Mapping Theorem:

Theorem 1 (Unified Covector Mapping Theorem) *Let $\mathbb{M}^{\lambda N}(\chi) \neq \emptyset$ and $\left\{ \widehat{\boldsymbol{\nu}}_0, [\widehat{\boldsymbol{\mu}}_k], [\widehat{\boldsymbol{\lambda}}_k] \right\} \in \widehat{\mathbb{M}}^{N\lambda}(\chi)$; then, the bijection, $\widehat{\mathbb{M}}^{N\lambda}(\chi) \sim \mathbb{M}^{\lambda N}(\chi)$, is given by,*

$$\boldsymbol{\lambda}^N(t_k) = \widehat{\boldsymbol{\lambda}}_k \quad \boldsymbol{\mu}^N(t_k) = \widehat{\boldsymbol{\mu}}_k, \quad \boldsymbol{\nu}_0 = \widehat{\boldsymbol{\nu}}_0 \tag{29}$$

Remark 2 *The statement of Theorem 1 is identical to that of Ref. [50] and is made possible through the construction of weighted interpolation. The weighted interpolants lead to a consistent pair of differentiation matrices, D and D^*, that are dual to each other. In this context, the LGL case turns out to be the special situation where the formal adjoint of D is the same as $-D$.*

Remark 3 *Note that Theorem 1 neither states $\boldsymbol{\lambda}^N(t_k) = \widetilde{\lambda}_k$, $\boldsymbol{\mu}^N(t_k) = \widetilde{\boldsymbol{\mu}}_k$, $\boldsymbol{\nu}_0 = \widetilde{\nu}_0$ nor does it imply $\boldsymbol{\lambda}^N(t_k) \approx \widetilde{\lambda}_k$, $\boldsymbol{\mu}^N(t_k) \approx \widetilde{\boldsymbol{\mu}}_k$, $\boldsymbol{\nu}_0 \approx \widetilde{\nu}_0$ as is sometimes erroneously interpreted. Furthermore, note that Eq. (29) is an <u>exact</u> relationship.*

Remark 4 *A simple counter example is constructed in Ref. [33] to show that a solution to Problem $B^{\lambda N}$ may not exist for Euler discretization no matter how small the mesh. This well-known phenomenon requires a proper technical modification to Theorem 1 similar to Polak's theory of consistent approximations. This aspect of Theorem 1 is rigorously proved in Refs. [30] and [28] for unit weight functions. The extension of this rigor to non-unit weight function is straightforward but lengthy. Thus, the main assumption in Theorem 1 that the multiplier set, $\mathbb{M}^{\lambda N}(\chi)$, is non empty can be eliminated. In other words, Theorem 1 remains valid under much weaker assumptions.*

Remark 5 *Because the set $\mathbb{M}^{N\lambda}(\chi)$ is "larger" (i.e. has more elements) than $\mathbb{M}^{\lambda N}(\chi)$, it de-sensitizes the sensitivities associated with solving Problem B^N. This fact is exploited in Ref. [52] to design a "guess-free" algorithm for solving Problem B.*

B. Some Remarks on the Covector Mapping Principle

The Covector Mapping Theorem was generated by an application of the Covector Mapping Principle (see Fig. 7). Since the introduction of Pontryagin's Principle, it has been known that the Minimum Principle may fail in the discrete-time domain if it is applied in exactly the same manner as in the continuous-time domain. For the Minimum Principle to hold exactly in the discrete-time domain, additional assumptions of convexity are required whereas no such assumptions are necessary for the continuous-time versions. This is because the continuous-time problem has a *hidden convexity*[49] that is implicitly used in proving theorems. Thus, when a continuous-time optimal control problem is discretized, this hidden convexity is not carried over to the discrete time domain resulting in a loss of information. If this information loss is not restored, the discrete-time solution may be spurious, not converge to the correct solution, or may even provide completely false results. A historical account of these issues, along with a simple counter example, is described in Ref. [33]. Further details are provided in the references contained in Ref. [33]. A thorough discussion of these issues and their relationship to advance concepts in optimal control theory has been developed by Mordukhovich.[48, 49]

The closure conditions introduced by Ross and Fahroo[26] are a form of matching conditions that are similar in spirit to Mordukhovich's matching conditions for Euler approximations. These conditions can be in primal space alone,[48, 51] or in primal-dual space.[34] Note however that the primal space conditions[48, 51] are obtained through dual space considerations. These new ideas reveal that dual space issues cannot be ignored even in the so-called "direct methods," for optimal control.

C. Software

A proper implementation of a PS method requires addressing all the numerical stability and accuracy issues and a case-by case approach to problem solving using PS approximations is inadvisable as it is tantamount to using first principles in every situation. Consequently, it is preferable for all the intricacies of a PS method to be implemented just once for a general problem in a form of a reusable software package. This exercise has been carried out in both OTIS and DIDO.[23] OTIS is in FORTRAN while DIDO is in MATLAB. While OTIS uses PS techniques as one of its many option, DIDO is exclusively based on PS methods alone. In addition, DIDO implements the spectral algorithm[53] to complete the "circuit" shown in Fig. 7. The generalized approach is implemented in an α-version of the software package. This software was used in the following sections that illustrate the principles and practice.

VI. Illustrating Some Key Points of the Main Theorem

In this section we will illustrate the many elements of our unified theory. From the statements of Theorem 1, it is clear that the LGL PS method is the most general of the Legendre PS methods and is hence applicable to all finite-horizon optimal control problems. Furthermore, note that although the Covector Mapping Theorem is meaningful only under convergence of the discretization, convergence issues are separate and distinct concepts. The apparently many nuances of the Unified Covector Mapping Theorem can be best illustrated by examples and counter examples.

American Institute of Aeronautics and Astronautics

A. Convergence of the Costates Does Not Imply Convergence of the Control

Obviously, the key variable in an optimal control problem is the control variable. Thus, convergence of the control variable implies convergence of the states and costates under fairly mild conditions; however, the opposite is not true.[54] To illustrate this point we once again consider Problem P_1. By a direct application of the Minimum Principle, it can be shown that the exact values of the costates are given by,

$$\lambda_x(t) \quad = \quad -2 \tag{30}$$
$$\lambda_v(t) \quad = \quad -1 \tag{31}$$

Figure 3 shows that the LGL controls have converged; thus, as expected, the LGL costates converge as illustrated in Fig. 8. On the other hand, the controls for the LG and LGR PS methods do not converge (see

Figure 8. Exact and LGL PS costates for Problem P_1.

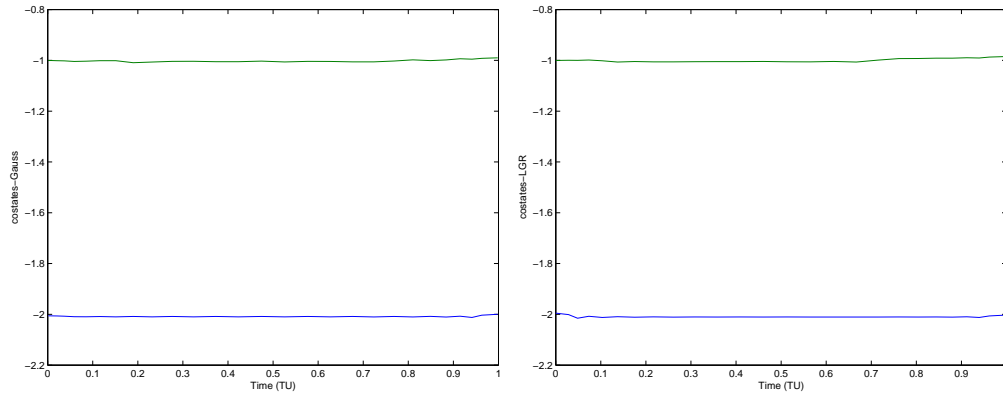Fig. 5) despite that the costates do indeed converge as illustrated in Fig. 9.

Figure 9. Exact and PS costates for Problem P_1: LG and LGR PS methods; compare Fig. 5.

B. Illustrating the Utility of the Covector Mapping Theorem

The premise of computational methods is that, unlike Problem P_1 most optimal control problems do not have analytical solutions. Consequently, any computational solution requires theoretical justification along with multiple and independent means to verify feasibility and optimality of the solution. The Covector Mapping Theorem, when applied correctly, provides one such test. To illustrate this point, we consider the much-studied orbit transfer problem of Moyer and Pinkham.[55] This widely-studied[31, 51, 56] classical problem,

American Institute of Aeronautics and Astronautics

has no analytical solution, and is given by,

$$\mathbf{x} = (r, \theta, v_r, v_t) \in \mathbb{R}^4, \quad \mathbf{u} = \beta \in \mathbb{R}$$

$$(P_2) \begin{cases} \text{Minimize} & J[\mathbf{x}(\cdot), \mathbf{u}(\cdot)] = -r(t_f) \\ \text{Subject to} & \dot{r}(t) = v_r(t) \\ & \dot{\theta}(t) = \frac{v_\theta}{r_2} \\ & \dot{v}_r(t) = \frac{v_\theta^2}{r} - \frac{1}{r^2(t)} + A(t)\sin\beta(t) \\ & \dot{v}_\theta(t) = -\frac{v_r(t)v_\theta(t)}{r(t)} + A(t)\cos\beta(t) \\ & t_f = 3.32 \\ & (r(0), \theta(0)) = (1, 0) \\ & (v_r(0), v_\theta(0)) = (0, 1) \\ & (v_r(t_f), v_\theta(t_f)) = (0, \sqrt{1/r(t_f)}) \end{cases}$$

where the state variables are the radial distance r, the angular distance θ, the radial velocity component v_r, and the transverse velocity component v_θ. The control variable is the thrust steering angle measured from the local horizontal β, and $A(t)$ is the continuous acceleration parameterized by,

$$A(t) = \frac{T}{(m_0 - |\dot{m}|t)} \tag{32}$$

where $T = 0.1405$, is the continuous thrust, $m_0 = 1.0$ is the initial mass, and $|\dot{m}| = 0.0749$ is the constant fuel consumption rate.

The state and control solution to this problem using the LGL PS method is shown in Fig. 10. Although,
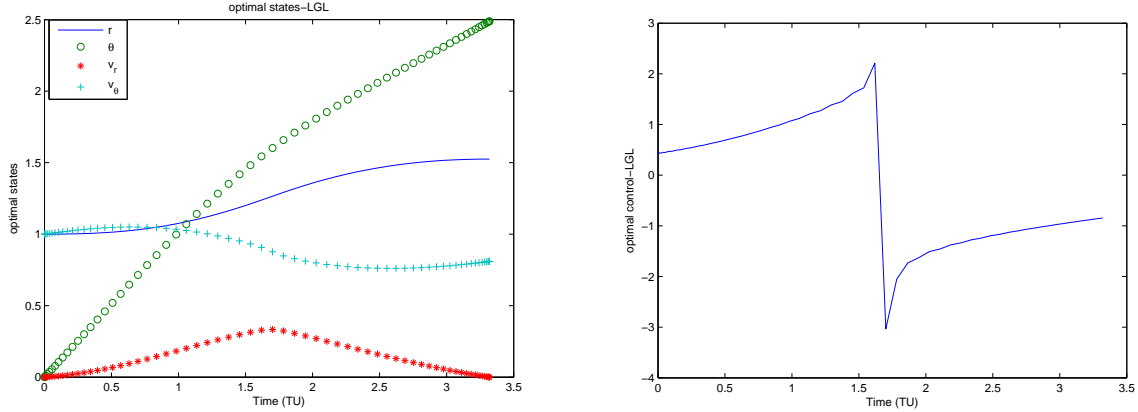

Figure 10. State and control trajectories for the LGL PS method.

this solution is identical to the validated one we obtained in Ref. [57] a decade ago, one of the advantages of the Covector Mapping Theorem is that it provides an independent means to test the optimality of the solution by an application of Pontrygain's Principle. To demonstrate this point, we begin by developing the necessary conditions for Problem P_2. It is straightforward to show that the adjoint equation for λ_θ along with the transversality condition indicates that,

$$\lambda_\theta(t) = 0 \tag{33}$$

The costates obtained by an application of the Covector Mapping Theorem are shown in Fig. 11. Clearly, Fig. 11 shows that $\lambda_\theta(t) = 0$ to within numerical errors. Thus, it is clear that the Covector Mapping Theorem is extremely useful in verifying the optimality of the candidate solution by a direct application of Pontryagin's Principle.
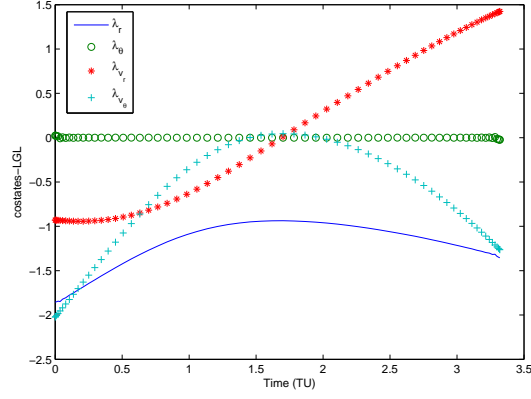
American Institute of Aeronautics and Astronautics

Figure 11. Costates obtained from the Covector Mapping Theorem; note that $t \mapsto \lambda_\theta = 0$

C. Illustrating The Multiplier Sets of Theorem 1

The concept of multiplier sets used in the statement of Theorem 1 is based on our prior work presented in Ref. [50]. At this point it is useful to illustrate how our multiplier sets can be used while drawing attention to their possible misuse as well.

It can be shown[29] that for Problem P_2, the map, $\chi \rightrightarrows \mathbb{M}^{N\lambda}(\chi)$, is indeed multivalued, and hence the notation, $\rightrightarrows$. Let $\widetilde{\Lambda} \in \mathbb{M}^{N\lambda}(\chi)$. Fig. 12 shows the multiplier trajectories obtained by one selection of the
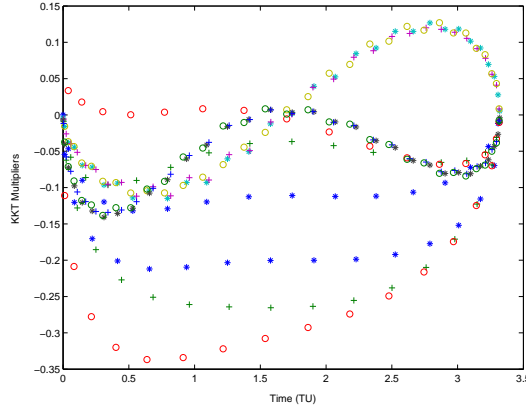

Figure 12. KKT multiplier trajectories, $t \mapsto \widetilde{\lambda}$, in the wrong Hilbert space and wrong selection.

set-valued map, $\mathbb{M}^{N\lambda}(\chi)$. This selection is based on a least-squares solution (LSQ) of minimizing $\left\| \widetilde{\Lambda} \right\|_2$,

$$(LSQ) \begin{cases} \text{Minimize} & \left\| \widetilde{\Lambda} \right\|_2^2 \\ \text{Subject to} & \widetilde{\Lambda} \in \mathbb{M}^{N\lambda}(\chi) \end{cases}$$

Comparing Figs. 11 and 12, it is clear that the multiplier trajectories of Fig. 12 are not the discrete costates; they are not even approximations as noted in Remark 3. Because the dual space is a weighted Hilbert space, one could credibly argue that the multiplier trajectories of Fig. 12 are in the wrong space. By selecting the multipliers in the correct Hilbert space but still based on the selection given by Problem LSQ, we get the multiplier trajectories shown in Fig. 13. Clearly, none of the KKT multipliers can be construed as discrete costate trajectories. The correct selection of multiplier trajectories are indeed obtained from Eq. (28). Not surprisingly, a proper use of the theory generates the right answer!
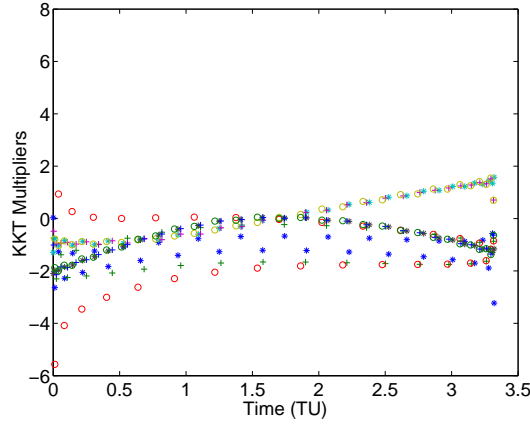
American Institute of Aeronautics and Astronautics

Figure 13. KKT multiplier trajectories, $t \mapsto \widetilde{\lambda}$, in the correct Hilbert space but wrong selection.

Note that it is incorrect to interpret any of the KKT multiplier trajectories as "wiggles" about the correct answer. Substantial details may be found in Refs. [25, 26, 28, 32–35, 40, 58] with various levels of sophistication and rigor.

Based on the primal space arguments put forth in Sec V.B, it is clear that the LGR and LG PS methods are inappropriate for Problem P_2 as well. Nonetheless, it is instructive to explore the consequences of selecting the wrong PS method, and examine its dual space consequences.

D. Illustrating the Consequences of an Improper Choice of a PS Method

From Theorem 1, it is clear that we should not use the LGR or LG PS methods for this problem. Regardless, it is instructive to examine their numerical performance in order to determine the practical consequences of choosing the wrong PS method. As already noted previously, what matters in optimal control problems is the convergence of the control variables; hence, we focus our attention on the convergence of the control.

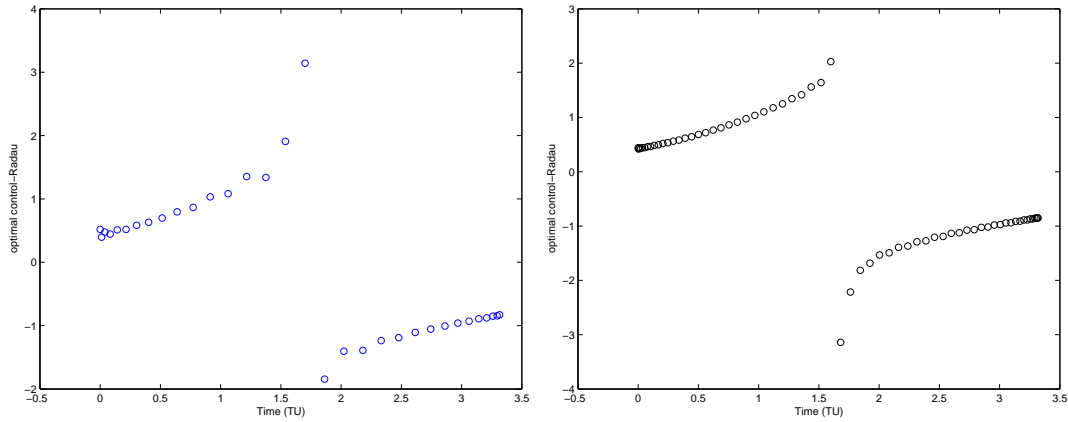The optimal controls for the LGR PS method is shown in Fig. 14. From this figure it is apparent

Figure 14. LGR controls for N=32 and n=64.

that the LGR PS method indicates convergence but at a slower rate than the LGL PS method. On the other hand, it is clear from Fig. 15. that the LG controls perform disastrously as the number of nodes are increased indicating the absence of convergence. As neither the LG nor the LGR PS methods are the correct PS methods for this problem, the numerical experimentation reveals that the wrong PS method does not necessarily generate the wrong answer, but that it is impossible to provide assurances on the validity of a solution obtained by the wrong method.
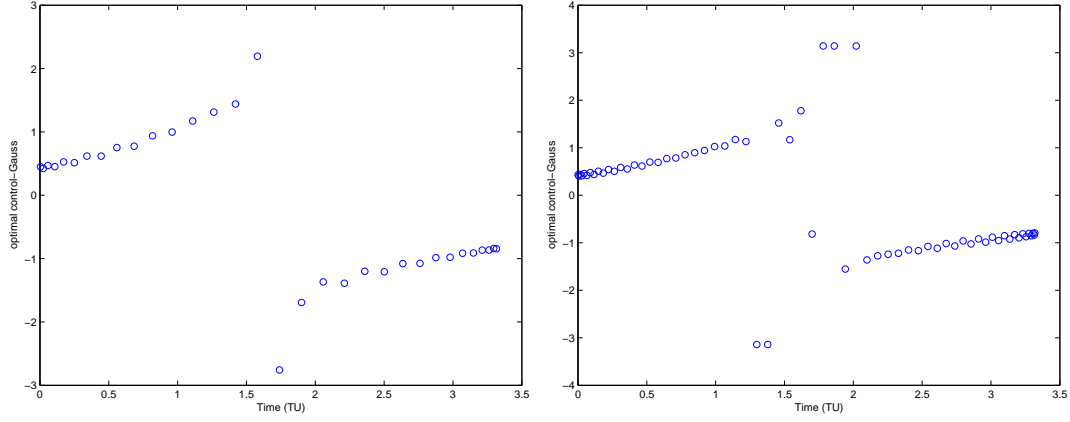
American Institute of Aeronautics and Astronautics

Figure 15. LG controls for N=32 and n=64. Note the divergence of the controls.

E. Some Remarks on the Completeness of Our Framework

Because the intent of this paper is on generalizing and unifying PS methods, we have chosen to describe our core ideas in terms of certain simplifications. For instance, we have chosen a distilled optimal control Problem B to describe the unified principles. Our choice of the distilled Problem B is not to be confused with the limitation of the ideas; rather, with additional bookkeeping all our ideas transfer trivially to substantially more complex problems than the one posed as Problem B. In the same spirit, we have restricted the discussions in this paper to Legendre-based PS methods. Our ideas apply equally to other PS methods as well. The key new concept we have proposed in this paper is contained in Eq. (18). *This is the notion of consistency in primal and dual functions for interpolation.* That is, the weight functions for primal interpolation is not necessarily the same as the weight function for dual interpolation. This notion of consistency is different from Polak's theory of consistent approximations. Thus PS methods for optimal control are now firmly established as being distinct from PS methods in other applications. Roughly speaking, PS methods for optimal control can be described as follows: For any chosen primal interpolant with primal weight function, $W(t)$, it is necessary to choose a consistent dual interpolant with weight function, $W^*(t)$. The weight function pair, $\{W(t), W^*(t)\}$, must be selected in a manner that generates the correct Hilbert space for the approximation of functions. Under the assumption of existence of a continuous-time solution in an appropriate Sobolev space, global convergence of the solution based on the consistency of the approximation can then be rigorously proved under mild and checkable conditions. Refs. [28, 30] present some of the foundations for such results. These theoretical developments explain the widely-demonstrated superior performance of PS methods for optimal control. A generalized presentation of these ideas is beyond the scope of this paper, but it is now clear why the Lobatto family of Legendre PS methods have a weight function of unity. For non-unit weight functions, the starting point for designing PS methods is Eq. (18).

VII. Conclusions

In recent years, PS methods have migrated from theory and computation to ground experiments as well as flight implementation. The high reliance on PS methods by industry demands that sound practices in computational verification and validation (V&V) be adopted. In sharp contrast to its theory in fluid dynamics, the proper framework for PS methods for optimal control is a weighted pre-Hilbert space. This unification of the apparently disparate PS methods reveals two major conclusions: (i) for finite-horizon optimal control problems, the correct PS methods are based on Gauss-Lobatto points while, (ii) for infinite-horizon optimal control problems, Gauss-Radau points form the right choice for discretization. The choice of Gauss points alone leads to information loss, particularly at the critical initial point because pure Gauss points are all interior to the interval. Nonetheless, Gauss points may be chosen for special finite-horizon problems such as when the boundary conditions are homogeneous. If the wrong PS method is chosen to solve the problem, numerical experiments show that it is possible to obtain seemingly valid results for

American Institute of Aeronautics and Astronautics

many problems, particularly for low orders of discretization. Because numerical solutions to optimal control problems are sought when no analytic solutions are available, suspicion may be cast on the validity of the solutions obtained by the wrong PS method. Industrial-strength finite-horizon optimal control problems have complex boundary conditions; consequently, the proper choice of PS methods for such problems are selectable from the Lobatto family. Because the Lobatto-based Legendre PS method is consistent with a unit weight function, the discovery of primal-dual weighted polynomials remained shielded until recently. Finally, when the correct PS method is used, the multiplier sets may be effectively used to de-sensitize any algorithmic sensitivities leading to a guess-free approach for solving the optimal control problem.

VIII.   Acknowledgments

References

[1] Kang, W. and Bedrossian, N., "Pseudospectral Optimal Control Theory Makes Debut Flight – Saves NASA $1M in Under 3 hrs," *SIAM News*, Vol. 40, No. 7, September 2007, Page 1.

[2] Sekhavat, P., Fleming A., and Ross, I. M., "Time-Optimal Nonlinear Feedback Control for the NPSAT1 Spacecraft," *Proceedings of the 2005 IEEE/ASME International Conference on Advanced Intelligent Mechatronics*, AIM 2005, 24-28 July 2005 Monterey, CA.

[3] Ross, I. M., Sekhavat, P., Fleming, A., and Gong, Q., "Pseudospectral Feedback Control: Foundations, Examples and Experimental Results," *Proceedings of the AIAA Guidance, Navigation and Control Conference*, Keystone, CO, August 2006. AIAA-2006-6354.

[4] Bedrossian, N., Bhatt, S., Lammers M., Nguyen, L., Zhang, Y., "First Ever Flight Demonstration of Zero Propellant Maneuver Attitude Control Concept," AIAA-2007-6734.

[5] Bedrossian, N., Bhatt, S., Lammers M., Nguyen, L., "Zero Propellant Maneuver: Flight Results for 180° ISS Rotation," *20th International Symposium on Space Flight Dynamics*, September 24-28, 2007, Annapolis, MD, NASA/CP-2007-214158.

[6] Infeld, S. I., "Optimization of Mission Design For Constrained Libration Point Missions," Ph.D. Dissertation, Department of Aeronautics and Astronautics, Stanford University, December 2005.

[7] Lu, P., Sun H., and Tsai, B., "Closed-Loop Endoatmospheric Ascent Guidance," *Journal of Guidance, Control and Dynamics*, Vol. 26, No. 2, pp.283-294, 2003.

[8] Melton, R. G., "Comparision of Direct Optimization Mehods Applied to Solar Sails," *AIAA/AAS Astrodynamics Specialists Conference and Exhibit*, 5-8 August 2002, Monterey, CA, AIAA 2002-4728.

[9] Paris, S. W., Riehl, J. P. and Sjaw, W. K., "Enhanced Procedures for Direct Trajectory Optimization Using Nonlinear Programming and Implicit Integration," AIAA/AAS Astodynamics Specialist Conference and Exhibit, 21-24 August 2006, Keystone, CO, AIAA 2006-6309.

[10] Riehl, J. P., Paris, S. W., and Sjaw, W. K., "Comparision of Implicit Integration Methods for Solving Aerospace Trajectory Optimization Problems," AIAA/AAS Astodynamics Specialist Conference and Exhibit, 21-24 August 2006, Keystone, CO, AIAA 2006-6033.

[11] Harada, M., and Bollino, K., "Optimal Trajectory of a Glider in Ground Effect and Wind Shear," *AIAA Guidance, Navigation and Control Conference*, San Francisco, CA, August 15-18, 2005. AIAA 2005-6474.

[12] Harada, M., Bollino, K., and Ross, I. M., "Minimum-Fuel Circling for an Unmanned Aerial Vehicle," *2005 JSASS-KSAS Joint International Symposium on Aerospace Engineering*, Nagoya, Japan, October 12-15, 2005.

[13] Hawkins, A. M., Fill, T. R., Proulx, R. J., Feron, E. M., "Constrained Trajectory Optimization for Lunar Landing," *AAS Spaceflight Mechanics Meeting*, Tampa, FL, January 2006, AAS 06-153.

[14] Infeld, S. I. and Murray, W., "Optimization of Stationkeeping for a Libration Point Mission," *AAS Spaceflight Mechanics Meeting*, Maui, HI, February 2004. AAS 04-150.

[15] J. Pietz and N. Bedrossian, "Moemtum Dumping Using Only CMGs," *Proceedings of the AIAA GNC Conference*, Austin, TX, 2003.

[16] Rea, J., "Launch Vehicle Trajectory Optimization Using a Legendre Pseudospectral Method," *Proceedings of the AIAA Guidance, Navigation and Control Conference*, Austin, TX, August 2003. Paper No. AIAA 2003-5640.

[17] S. Stanton, R. Proulx and C. D'Souza, Optimal orbit transfer using a Legendre pseudospectral method, *AAS/AIAA Astrodynamics Specialist Conference*, AAS-03-574, Big Sky, MT, August 3-7, 2003.

[18] P. Williams, Application of Pseudospectral Methods for Receding Horizon Control, *J. of Guid., Contr. and Dyn.*, Vol.27, No.2., 2004, pp.310-314.

[19] P. Williams, C. Blanksby and P. Trivailo, Receding horizon control of tether system using quasilinearization and Chebyshev pseudospectral approximations, *AAS/AIAA Astrodynamics Specialist Conference*, Big Sky, MT, August 3-7, 2003, Paper AAS 03-535.

[20] H. Yan and K. T. Alfriend, Three-axis Magnetic Attitude Control Using Pseudospectral Control Law in Eccentric Orbits, *AAS Spaceflight Mechanics Meeting*, Tampa, FL, January 2006, AAS 06-103.

[21] Fahroo, F. and Ross, I. M., "On Discrete-Time Optimality Conditions for Pseudospectral Methods," *Proceedings of the AIAA/AAS Astrodynamics Conference*, Keystone, CO, August 2006. AIAA-2006-6304

[22]Paris, S. W. and Hargraves, C. R., *OTIS 3.0 Manual*, Boeing Space and Defense Group, Seattle, WA, 1996.

[23]Ross, I. M., *A Beginner's Guide to DIDO: A MATLAB application package for solving optimal control problems*, Elissar, Monterey, CA, 2007.

[24]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.

[25]Ross, I. M. and Fahroo, F., "A Pseudospectral Transformation of the Covectors of Optimal Control Systems," *Proceedings of the First IFAC Symposium on System Structure and Control*, Prague, Czech Republic, 29-31 August 2001.

[26]Ross, I. M. and Fahroo, F., "Legendre Pseudospectral Approximations of Optimal Control Problems," *Lecture Notes in Control and Information Sciences*, Vol.295, Springer-Verlag, New York, 2003.

[27]Ross, I. M. and Fahroo, F., "A Perspective on Methods for Trajectory Optimization," *Proceedings of the AIAA/AAS Astrodynamics Conference*, Monterey, CA, August 2002. AIAA Paper No. 2002-4727.

[28]Gong, Q., Ross, I. M., Kang, W. and Fahroo, F., "On the Pseudospectral Covector Mapping Theorem for Nonlinear Optimal Control," to appear in *Proceedings of the 45th IEEE Conference on Decision and Control*, San Diego, CA, December 13-15, 2006.

[29]Gong, Q., Ross, I. M., Kang, W., and Fahroo, F., "Connections Between the Covector Mapping Theorem and Convergence of Pseudospectral Methods for Optimal Control," *to appear in Computational Optimization and Applications*, 2007.

[30]Gong, Q., Kang, W. and Ross, I. M., "A Pseudospectral Method for the Optimal Control of Constrained Feedback Linearizable Systems," *IEEE Transactions on Automatic Control*, Vol. 51, No. 7, July 2006, pp. 1115-1129.

[31]Vinter, R. B., *Optimal Control*, Birkhäuser, Boston, MA, 2000.

[32]Ross, I. M., "A Framework for Analyzing Discrete Approximations in Optimal Control," *SIAM Conference on Mathematics for Industry: Challenges and Frontiers*, Invited Talk, Toronto, Ontario, Canada, 13-15 October 2003.

[33]Ross, I. M., "A Historical Introduction to the Covector Mapping Principle," *Advances in the Astronautical Sciences*, Vol. 123, Univelt, San Diego, CA 2006, pp. 1257-1278.

[34]Ross, I. M., "A Roadmap for Optimal Control: The Right Way to Commute," *Annals of the New York Academy of Sciences*, Vol. 1065, December 2005, pp. 210-231.

[35]Ross, I. M., "Certain Connections in Optimal Control Theory and Computation," *Proceedings of the Symposium on New Trends in Nonlinear Dynamics and Control and Their Applications*, Invited Talk, 18-19 October 2002, Monterey, CA.

[36]Davis, P., and Rabinowitz, P., *Methods of Numerical Integraion,* Academic Press, 1975.

[37]Karniadakis, G., and Sherwin, S., *Spectal/hp Element Methods for Computational Fluid Dynamics,* Oxford University Press, Oxford, 2005.

[38]Fahroo, F. and Ross, I. M., "Pseudospectral Methods for Infinite Horizon Nonlinear Optimal Control Problems," *AIAA Guidance, Navigation and Control Conference*, San Francisco, CA, 2005.

[39]Ross, I. M., Gong, Q., Fahroo, F. and Kang, W., "Practical Stabilization Through Real-Time Optimal Control," *2006 American Control Conference*, Minneapolis, MN, June 14-16 2006.

[40]Ross, I. M. and Fahroo, F., "Discrete Verification of Necessary Conditions for Switched Nonlinear Optimal Control Systems," *Proceedings of the American Control Conference*, June 2004, Boston, MA.

[41]Fahroo, F. and Ross, I. M., "Pseudospectral Methods for Infinite Horizon Optimal Control Problems," *Journal of Guidance, Control and Dynamics*, July-August, 2008.

[42]Trefethen, L. N., *Spectral Methods in MATLAB,* SIAM, Philadelphia, PA, 2000.

[43]Hager, W. W., "Numerical Analysis in Optimal Control," *International Series of Numererical Mathematics*, Hoffmann, K.-H. Lasiecka, I., Leugering, G., Sprekels, J., and Troeltzsch, F., Eds., Birkhäuser, Basel, Switzerland, 2001, Vol. 139, pp. 83–93.

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

[45]Weideman, J.A.C., "Spectral Methods Based on Non-Classical Orthogonal Polynomials," *The Proceedings of the Oberwolfach Meeting on Application and Computation of Orthogonal Polynomials*, March 1998.

[46]J. T. Betts, *Practical Methods for Optimal Control Using Nonlinear Programming*, SIAM, Philadelphia, PA, 2001.

[47]Ross, I. M., Gong, Q. and Sekhavat, P., "Low-Thrust, High-Accuracy Trajectory Optimization," *Journal of Guidance Control and Dynamics*, Vol. 30, No. 4, 2007, pp. 921-933

[48]Mordukhovich, B. S., *Variational Analysis and Generalized Differentiation, I: Basic Theory*, vol. 330 of Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences] Series, Springer, Berlin, 2005.

[49]Mordukhovich, B. S., *Variational Analysis and Generalized Differentiation, II: Applications*, vol. 331 of Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences] Series, Springer, Berlin, 2005.

[50]Ross, I. M. and Fahroo, F., "Discrete Verification of Necessary Conditions for Switched Nonlinear Optimal Control Systems," *Proceedings of the American Control Conference*, June 2004, Boston, MA.

[51]Hager, W. W., "Runge-Kutta Methods in Optimal Control and the Transformed Adjoint System," *Numerische Mathematik,* Vol. 87, 2000, pp. 247-282.

[52]Ross, I. M., and Gong, Q., "Guess-Free Trajectory Optimization," *AIAA/AAS Astrodynamics Specialist Conference and Exhibit*, Honolulu, Hawaii, August 18-21, 2008. AIAA Paper 2008-6273.

[53]Gong, Q., Fahroo, F. and Ross, I. M., A spectral algorithm for pseudospectral methods in optimal control, *AIAA Journal of Guidance, Control and Dynamics,* Vol. 31, No. 3, pp. 460-471, 2008.

[54]Fahroo, F. and Ross, I. M., "Convergence of the costates does not imply convergence of the control," *Journal of Guidance, Control and Dynamics*, to appear.

[55]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, New York, Academic Press, 1964, pp. 91-105.

[56]Fahroo, F. and Ross, I. M., "Costate Estimation by a Legendre Pseudospectral Method," *Journal of Guidance, Control and Dynamics*, Vol.24, No.2, March-April 2001, pp.270-277.

[57]Fahroo, F. and Ross, I. M., "Costate Estimation by a Legendre Pseudospectral Method," *AIAA Guidance, Navigation and Control Conference*, 1998.

[58]Gong, Q., Ross, I. M., Kang, W. and Fahroo, F., "Dual Convergence of the Legendre Pseudospectral Method for Solving Nonlinear Constrained Optimal Control Problems," *Proceedings of the IASTED International Conference on Intelligent Systems and Control*, Cambridge, MA, 2005.

American Institute of Aeronautics and Astronautics