Proceedings of the 45th IEEE Conference on Decision & Control
Manchester Grand Hyatt Hotel
San Diego, CA, USA, December 13-15, 2006

On the Pseudospectral Covector Mapping Theorem for Nonlinear Optimal Control

Qi Gong, I. Michael Ross, Wei Kang and Fariba Fahroo

Abstract— **In recent years, a large number of nonlinear optimal control problems have been solved by pseudospectral (PS) methods. In an effort to better understand the PS approach to solving control problems, we present convergence results for problems with mixed state and control constraints. A set of sufficient conditions are proved under which the solution of the discretized optimal control problem converges to the continuous solution. Conditions for the convergence of the duals are described and illustrated. This leads to a clarification of Covector Mapping Theorem and its connections to constraint qualifications.**

I. INTRODUCTION

The main difficulties in solving a state- and control-constrained nonlinear optimal control problem are in seeking a closed-form solution to the Hamilton-Jacobi equations, or in solving the canonical Hamiltonian equations resulting from an application of the Minimum Principle. Over the last decade, a third alternative based on discrete approximations has gained wide popularity [2], [4], [5], [10], [11], [13], [14], [16] as a result of significant progress in computation and theory. In simple terms, the idea can be characterized as discretizing the optimal control problem and solving the resulting large-scale finite-dimensional optimization problem. The simplicity of this approach belies a wide range of deep theoretical issues (see [16]) that lie at the intersection of approximation theory, control theory and optimization. Regardless, a wide variety of industrial-strength optimal control problems have been solved by this approach [2], [13], [15], [19], [21], [27].

In this paper we focus on pseudospectral (PS) methods. PS methods were largely developed in the 1970s for solving partial differential equations arising in fluid dynamics and meteorology [3], and quickly became "one of the big three technologies for the numerical solution of PDEs" [28]. During the 1990s, PS methods were introduced for solving optimal control problems; and since then, have gained considerable attention [5], [6], [13], [15], [23], [27], [29], [30], particularly in solving aerospace control problems. Examples range from lunar guidance [13], magnetic control [30], orbit transfers [27], tether libration control [29], ascent guidance [15] and a host of other problems. As a result of its considerable success that includes experimental validation [26], NASA's next generation of the OTIS software package [18] incorporates the Legendre PS method as a problem solving option. Further details on NASA's plans are described at: http:// trajectory.grc.nasa.gov/projects/lowthrust.shtml. In addition, the commercially available software package, DIDO [22], exclusively uses PS methods for solving optimal control problems.

Because PS methods are of recent vintage, when compared to, say, Runge-Kutta (RK) methods, a theory for PS discretizations is an emerging area of interest. In recent years, it has become clear that standard convergence theorems frequently employed in the analysis of differential equations are not applicable to discretizations of optimal control problems. For example, Hager [11] has shown that "convergent" RK methods can diverge while Betts et al [1] demonstrate that "nonconvergent" RK methods can converge. Furthermore, with regards to PS methods, its marked differences with other methods implies that a new approach is needed to address some fundamental questions. In this paper, we address some of these basic questions. For example, does the discretized problem always have a solution if a solution to the continuous-time problems exists? If so, under what conditions? Does the discretized solution converge to the continuous optimal solution? These questions are of interest not only from a theoretical standpoint, but are also of great practical value, particularly in the real-time computation of optimal control [25], [26].

In this paper, we strengthen earlier results and weaken prior assumptions. For example, in [10] the existence and the convergence results of PS methods are proved for nonlinear systems in feedback linearizable normal form. In this paper, we extend these results to the general nonlinear systems, and show that the discrete dynamics must be relaxed to guarantee feasibility. In [24] a set of "closure conditions" were identified to map the Karush-Kuhn-Tucker (KKT) multipliers associated with the discretized optimal control problem to the dual variables associated with the continuous-time optimal control problem. Unlike Hager's RK method which imposes additional conditions on the primal problem (i.e. coefficients of the integration scheme), the conditions of [24] imposes constraints on both the primal and dual variables. In the absence of a convergence theorem, this procedure requires solving a difficult primal-dual mixed complementarity problem (MCP). In this paper, we prove that for constrained optimal control problems, the solution of the discretized optimal control problem converges to the solution of the

This research was supported in part by SAF and AFOSR under grant F1ATA0-60-6-2G002.

Dept. of Elec. & Comp. Engr., Univ. of Texas at San Antonio, qi.gong@utsa.edu

Dept. of Mech. and Astro. Engr., Naval Postgraduate School, Monterey, CA, 93943, imross@nps.edu

Dept. of Applied Math., Naval Postgraduate School, Monterey, CA, 93943, wkang@nps.edu

Dept. of Applied Math., Naval Postgraduate School, Monterey, CA, 93943, ffahroo@nps.edu

1-4244-0171-2/06/$20.00 ©2006 IEEE.

continuous optimal control problem. Thus, the MCP may be replaced by simpler NLP techniques and solvers. More importantly, we demonstrate why the convergence of the primal variables does not necessarily imply the convergence of the KKT multipliers to the continuous costate. This leads to a clarification of the closure conditions of [24] that ensure the convergence of the duals. A simple example is introduced to tie these ideas to constraint qualifications.

Throughout the paper we make extensive use of Sobolev spaces [3], $W^{m,p}$, that consists of functions, $\xi : [-1,1] \to \mathbb{R}$ whose j-th distributional derivative, $\xi^{(j)}$, lies in L^p for all $0 \le j \le m$ with the norm,

$$\| \xi \|_{W^{m,p}} = \sum_{j=0}^{m} \| \xi^{(j)} \|_{L^p}$$

For notational ease, we suppress the dependence of $W^{m,p}$ on vector-valued functions.

II. The Problem and its Discretization

Problem B: Determine the state-control function pair, $t \mapsto (x,u) \in \mathbb{R}^{N_x} \times \mathbb{R}^{N_u}$, that minimize the cost function

$$J[x(\cdot),u(\cdot)] = \int_{-1}^{1} F(x(t),u(t))\,dt + E(x(-1),x(1))$$

subject to the dynamics,

$$\dot{x}(t) = f(x(t),u(t)) \tag{1}$$

endpoint conditions

$$e(x(-1),x(1)) = 0 \tag{2}$$

and path constraints

$$h(x(t),u(t)) \le 0 \tag{3}$$

It is assumed that $F : \mathbb{R}^{N_x} \times \mathbb{R}^{N_u} \to \mathbb{R}$, $E : \mathbb{R}^{N_x} \times \mathbb{R}^{N_x} \to \mathbb{R}$, $f : \mathbb{R}^{N_x} \times \mathbb{R}^{N_u} \to \mathbb{R}^{N_x}$, $e : \mathbb{R}^{N_x} \times \mathbb{R}^{N_x} \to \mathbb{R}^{N_e}$, and $h : \mathbb{R}^{N_x} \times \mathbb{R}^{N_u} \to \mathbb{R}^{N_h}$, are continuously differentiable with respect to their arguments and their gradients are Lipschitz continuous over the domain. In order to apply the first order necessary conditions, appropriate constraint qualifications are implicitly assumed throughout the rest of the paper. In addition to these standard assumptions, we assume that an optimal solution $(x^*(\cdot),u^*(\cdot))$ exists with the optimal state, $x^*(\cdot) \in W^{m,\infty},\ m \ge 2$. Note that, if $x^*(t)$ is C^1 and $\dot{x}^*(t)$ has bounded derivative everywhere except for finitely many points on the closed interval $t \in [-1,1]$, then $x^*(\cdot) \in W^{2,\infty}$. On the other hand, from Sobolev's Imbedding Theorems [3], any function $x^*(\cdot) \in W^{m,\infty},\ m \ge 2$ must have continuous $(m-1)$-th order classical derivatives on $[-1,1]$. Therefore, this condition requires the optimal state $x^*(t)$ be at least continuously differentiable.

In the Legendre PS approximation of Problem B, the basic idea is to approximate $x(t)$ by N-th order Lagrange polynomials $x^N(t)$ based on the interpolation at the Legendre-Gauss-Lobatto (LGL) quadrature nodes, i.e.

$$x(t) \approx x^N(t) = \sum_{k=0}^{N} x^N(t_k)\phi_k(t),$$

where t_k are LGL nodes defined as,

$$\begin{aligned} &t_0 = -1,\ \ t_N = 1 \\ &t_k,\ \text{for } k = 1,2,\ldots,N-1,\ \text{are the roots of } \dot{L}_N(t) \end{aligned}$$

where $\dot{L}_N(t)$ is the derivative of the N-th order Legendre polynomial, $L_N(t)$. The Lagrange interpolating polynomial $\phi_k(t)$ is defined by

$$\phi_k(t) = \frac{1}{N(N+1)L_N(t_k)} \frac{(t^2-1)\dot{L}_N(t)}{t-t_k}. \tag{4}$$

It is readily verifiable that $\phi_k(t_j) = 1$, if $k = j$ and $\phi_k(t_j) = 0$, if $k \ne j$. The derivative of the i-th state $x_i(t)$ at the LGL node t_k can be approximated by

$$\dot{x}_i(t_k) \approx \dot{x}_i^N(t_k) = \sum_{j=0}^{N} D_{kj} x_i^N(t_j), \qquad i = 1,2,\ldots,N_x$$

where $(N+1) \times (N+1)$ differentiation matrix D is defined as

$$D_{ik} = \begin{cases} \frac{L_N(t_i)}{L_N(t_k)} \frac{1}{t_i - t_k}, & \text{if } i \ne k; \\ -\frac{N(N+1)}{4}, & \text{if } i = k = 0; \\ \frac{N(N+1)}{4}, & \text{if } i = k = N; \\ 0, & \text{otherwise} \end{cases} \tag{5}$$

Let $\bar{x}_k = x^N(t_k)$, $k = 0,1,\ldots,N$. In a standard PS method, the continuous differential equation is approximated by the following nonlinear algebraic equations

$$\sum_{i=0}^{N} \bar{x}_i D_{ki} - f(\bar{x}_k,\bar{u}_k) = 0, \qquad k = 0,1,\ldots,N \tag{6}$$

where $\bar{u}_k$ is taken to be analogous to $\bar{x}_k$. This discretization is used in [5], [6], [24] for optimal control problems. It will be apparent shortly that a feasible solution to (6) may not exist; hence, to guarantee feasibility of the discretization, we propose the following relaxation,

$$\left\| \sum_{i=0}^{N} \bar{x}_i D_{ki} - f(\bar{x}_k,\bar{u}_k) \right\|_\infty \le (N-1)^{\frac{3}{2}-m}, \tag{7}$$

Deferring a justification of this relaxation, note that when N tends to infinity, the difference between conditions (6) and (7) vanishes, since m, by assumption, is greater than or equal to 2. Throughout the paper, we use the "bar" notation to denote discretized variables. Note that the subscript in $\bar{x}_k$ denotes an evaluation of the approximate state, $x^N(t) \in \mathbb{R}^{N_x}$, at the node t_k whereas $x_k(t)$ denotes the k-th component of the exact state. The endpoint conditions and constraints are approximated in a similar fashion

$$\|e(\bar{x}_0,\bar{x}_N)\|_\infty \le (N-1)^{\frac{3}{2}-m} \tag{8}$$

$$h(\bar{x}_k,\bar{u}_k) \le (N-1)^{\frac{3}{2}-m} \cdot \mathbf{1}, \quad k = 0,\ldots,N \tag{9}$$

where $\mathbf{1}$ denotes $[1,\ldots,1]^T$.

Finally, the cost functional $J[x(\cdot),u(\cdot)]$ is approximated by the Gauss-Lobatto integration rule,

$$J[x(\cdot),u(\cdot)] \approx \bar{J}^N(\bar{X},\bar{U}) = \sum_{k=0}^{N} F(\bar{x}_k,\bar{u}_k)w_k + E(\bar{x}_0,\bar{x}_N)$$

where w_k are the LGL weights given by

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

and $\bar{X} = [\bar{x}_0,\ldots,\bar{x}_N]$, $\bar{U} = [\bar{u}_0,\ldots,\bar{u}_N]$.

Since practical solutions are bounded, the following constraints are added

$$\{\bar{x}_k \in \mathbb{X},\ \bar{u}_k \in \mathbb{U},\ k = 0, 1, \ldots, N\}$$

where $\mathbb{X}$ and $\mathbb{U}$ are two compact sets representing the search region and containing the continuous optimal solution $(x^*(t), u^*(t))$. Hence, the optimal control Problem B is approximated by an NLP with $\bar{J}^N$ as the objective function and (7), (8) and (9) as constraints; this is summarized as:

Problem B^{N} **:** Find $\bar{x}_k \in \mathbb{X}$ and $\bar{u}_k \in \mathbb{U}$, $k = 0, 1, \ldots, N$, that minimize

$$\bar{J}^N(\bar{X}, \bar{U}) = \sum_{k=0}^{N} F(\bar{x}_k, \bar{u}_k) w_k + E(\bar{x}_0, \bar{x}_N) \tag{10}$$

subject to

$$\left\| \sum_{i=0}^{N} \bar{x}_i D_{ki} - f(\bar{x}_k, \bar{u}_k) \right\|_\infty \leq (N-1)^{\frac{3}{2}-m} \tag{11}$$

$$||e(\bar{x}_0, \bar{x}_N)||_\infty \leq (N-1)^{\frac{3}{2}-m} \tag{12}$$

$$h(\bar{x}_k, \bar{u}_k) \leq (N-1)^{\frac{3}{2}-m} \cdot \mathbf{1} \tag{13}$$

III. Feasibility of Problem B^{N}

In Eulerian discretizations, for any given initial condition and control discretization, the states are uniquely determined. Hence, there always exists a feasible solution to the discretized dynamic system. For RK methods, a similar property holds if the mesh is sufficiently dense [11]. For pseudospectral methods such an existence result for controlled differential equations is not readily apparent. There are two main issues. PS methods are fundamentally different than traditional methods (like Euler or RK) in that they focus on approximating the tangent bundle rather than the differential equation. Since the differential equation is imposed over discrete points, in standard PS methods the boundary conditions are typically handled by not imposing the differential equations over the boundary [3], [28]. This technique cannot be used for controlled differential equations as it implies that the control can take arbitrary values at the boundary. This is one of the many reasons why PS methods for control are different from their counterparts in other fields. A counter example in [10] shows that (6) may not have any feasible solution. In [10], the feasibility problem was circumvented by restricting the dynamics to be feedback linearizable. In this paper, we relax (6) to (11) so that even general nonlinear systems can be guaranteed a feasible solution to Problem B^{N} as proved in Theorem 1 below. First, we need the following lemma.

Lemma 1: [3] Given any function $\xi \in W^{m,\infty}$, $t \in [-1, 1]$, there is a polynomial $p^N(t)$ of degree N or less, such that

$$|\xi(t) - p^N(t)| \leq C C_0 N^{-m}, \quad \forall\ t \in [-1, 1]$$

where C is a constant independent of N and $C_0 = \|\xi\|_{W^{m,\infty}}$. ($p^N(t)$ with the smallest norm $||\xi(t) - p^N(t)||_\infty$ is called the N-th order best polynomial approximation of $\xi(t)$ in the norm of L^∞ .)

Theorem 1: Given any feasible solution, $t \mapsto (x, u)$, for Problem B, suppose $x(\cdot) \in W^{m,\infty}$ with $m \geq 2$. Then, there exists a positive integer N_1 such that, for any $N > N_1$, Problem B^{N} has a feasible solution, $(\bar{x}_k, \bar{u}_k)$. Furthermore, the feasible solution satisfies $\bar{u}_k = u(t_k)$ and

$$\|x(t_k) - \bar{x}_k\|_\infty \leq L(N-1)^{1-m}, \tag{14}$$

for all $k = 0, \ldots, N$, where t_k are LGL nodes and L is a positive constant independent of N.

Proof: Let $p(t)$ be the $(N-1)$-th order best polynomial approximation of $\dot{x}(t)$ in the norm of L^∞ . By Lemma 1 there is a constant C_1 independent of N such that

$$\|\dot{x}(t) - p(t)\|_\infty \leq C_1(N-1)^{1-m}, \quad \forall t \in [-1, 1] \tag{15}$$

Define

$$\begin{aligned} x^N(t) &= \int_{-1}^{t} p(\tau) d\tau + x(-1) \\ \bar{x}_k &= x^N(t_k) \\ \bar{u}_k &= u(t_k) \end{aligned} \tag{16}$$

From (15),

$$\|x(t) - x^N(t)\|_\infty \leq 2C_1(N-1)^{1-m}, \quad \forall t \in [-1, 1] \tag{17}$$

It follows that both $x(t_k)$ and $\bar{x}_k$ are contained in some compact set whose boundary is independent of N. On this compact set, because f is continuously differentiable, it must be Lipschitz continuous. By definition, $x^N(t)$ is a polynomial of degree less than or equal to N. It is known (see [3]) that, for any polynomial of degree less than or equal to N, its derivative at the LGL nodes $t_0, \ldots, t_N$ are exactly equal to the value of the polynomial at the nodes multiplied by the differential matrix D, which is defined by (5). Thus we have

$$\sum_{i=0}^{N} \bar{x}_i D_{ki} = \dot{x}^N(t_k) \tag{18}$$

Therefore,

$$\begin{aligned} \left\| \sum_{i=0}^{N} \bar{x}_i D_{ki} - f(\bar{x}_k, \bar{u}_k) \right\|_\infty &\leq \left\| \dot{x}^N(t_k) - \dot{x}(t_k) \right\|_\infty \\ &\quad + ||\dot{x}(t_k) - f(\bar{x}_k, \bar{u}_k)||_\infty \\ &\leq ||p(\tau_k) - \dot{x}(\tau_k)||_\infty + ||f(x(t_k), u(t_k)) - f(\bar{x}_k, \bar{u}_k)||_\infty \\ &\leq C_1(N-1)^{1-m} + C_2\|x(t_k) - x^N(t_k)\|_\infty \\ &\leq C_1(1 + 2C_2)(N-1)^{1-m} \end{aligned}$$

where C_2 is the Lipschitz constant of f with respect to x. Since there exists a positive integer N_1 such that, for all $N > N_1$,

$$C_1(1 + 2C_2)(N-1)^{1-m} \leq (N-1)^{\frac{3}{2}-m}$$

Hence, (11) holds for all $N > N_1$.

As for the constraint (13), because h is continuously differentiable, the following estimate holds.

$$\begin{aligned} &\|h(x(t), u(t)) - h(x^N(t), u(t))\|_\infty \\ &\quad \leq C_3\|x(t) - x^N(t)\|_\infty \leq 2C_1C_3(N-1)^{1-m} \end{aligned}$$

where C_3 is the Lipschitz constant of h with respect to x which is independent of N. Hence

$$\begin{aligned} h(\bar{x}_k, \bar{u}_k) &\leq h(x(t_k), u(t_k)) + 2C_1C_3(N-1)^{1-m} \cdot \mathbf{1} \\ &\leq 2C_1C_3(N-1)^{1-m} \cdot \mathbf{1} \end{aligned}$$

Thus, the constraint (13) holds for all $N \geq N_1$. As for the endpoint condition (12), it can be proved in a similar fashion.

Thus, we have constructed a feasible solution $(\bar{x}_k, \bar{u}_k)$ for Problem B^N. Finally, (14) follows directly from (17). ■

Remark 1: In practice, we use a small number, $\delta_P > 0$ as a feasibility tolerance. Then, Theorem 1 guarantees that for any δ_P howsoever small, (11)–(13) always has a solution provided a sufficiently large number of nodes are chosen. Furthermore, the right hand side of (11)–(13) converges to zero as N tends to infinity.

IV. Convergence of the Primal Variables

With an existence result in hand, we now establish the convergence of the primal variables, (x, u). That is, to show the existence of a sequence of optimal solutions of Problem B^N converging to an optimal solution of Problem B. The method we used is similar in spirit to Polak's theory of consistent approximations [20]. We show that, under certain conditions, the sequence of finite dimensional nonlinear programming, Problem B^N, consistently approximate the infinite dimensional continuous optimal control Problem B.

Let $(\bar{x}_k^*, \bar{u}_k^*)$, $k = 0, 1, \ldots, N$, be an optimal solution to Problem B^N. Let $x^N(t) \in \mathbb{R}^{N_x}$ be the N-th order interpolating polynomial of $(\bar{x}_0^*, \ldots, \bar{x}_N^*)$ and $u^N(t) \in \mathbb{R}^{N_u}$ be any interpolant of $(\bar{u}_0^*, \ldots, \bar{u}_N^*)$, i.e.

$$x^N(t) = \sum_{k=0}^{N} \bar{x}_k^* \phi_k(t), \qquad u^N(t) = \sum_{k=0}^{N} \bar{u}_k^* \psi_k(t)$$

where $\phi_k(t)$ is the Lagrange interpolating polynomial defined by (4) and $\psi_k(t)$ is any continuous function such that $\psi_k(t_j) = 1$, if $k = j$ and $\psi_k(t_j) = 0$, if $k \neq j$. Note that $u^N(t)$ is not necessarily a polynomial. Typically, we use linear or spline functions for interpolating $(\bar{u}_0^*, \ldots, \bar{u}_N^*)$. Now consider a sequence of Problems B^N with N increasing from N_1 to infinity. Correspondingly, we get a sequence of discrete optimal solutions $\{(\bar{x}_k^*, \bar{u}_k^*), k = 0, \ldots, N\}_{N=N_1}^{\infty}$ and their interpolating function sequence $\{x^N(t), u^N(t)\}_{N=N_1}^{\infty}$.

Definition 1: A continuous function $\rho(t)$ is called the uniform accumulation point of a function sequence $\{\rho^N(t)\}_{N=0}^{\infty}$, $t \in [-1, 1]$, if there is a subsequence of $\{\rho^N(t)\}_{N=0}^{\infty}$ that uniformly converges to $\rho(t)$.

Assumption 1: x_0^{∞} is an accumulation point of the first element (i.e. $k = 0$) of the sequence, $\{\bar{x}_k^*, k = 0, \ldots, N\}_{N=N_1}^{\infty}$.

Remark 2: In many optimal control problems, an initial value of the state is fixed by the endpoint condition. Then, from (12), it is easy to verify that Assumption 1 is automatically satisfied.

Lemma 2: [8] Let t_k, $k = 0, 1, \ldots, N$, be the LGL nodes, and w_k be the LGL weights. Suppose $\xi(t)$ is Riemann integrable; then,

$$\int_{-1}^{1} \xi(t) dt \quad = \quad \lim_{N \to \infty} \sum_{k=0}^{N} \xi(t_k) w_k$$

Theorem 2: Let $\{(\bar{x}_k^*, \bar{u}_k^*), 0 \leq k \leq N\}_{N=N_1}^{\infty}$ be a sequence of optimal solutions of Problem B^N satisfying Assumption 1, and $(x^N(t), u^N(t))_{N=N_1}^{\infty}$ be their interpolating function sequence. Let the pair of continuous functions, $(q(t), u^{\infty}(t))$, be any uniform accumulation point of the sequence $(\dot{x}^N(t), u^N(t))_{N=N_1}^{\infty}$. Then, $u^{\infty}(t)$ is an optimal control to the original continuous Problem B, and $x^{\infty}(t) = \int_{-1}^{t} q(\tau) d\tau + x_0^{\infty}$ is the corresponding optimal trajectory.

Proof: By definition, there is a subsequence $N_i \in 0, 1, \cdots$, with $\lim_{i \to \infty} N_i = \infty$, such that

$$\lim_{i \to \infty} (\dot{x}^{N_i}(t), u^{N_i}(t)) \quad = \quad (q(t), u^{\infty}(t)).$$

It is easy to show (under Assumption 1)

$$\lim_{i \to \infty} x^{N_i}(t) \quad = \quad x^{\infty}(t) \tag{19}$$

uniformly on $t \in [-1, 1]$. The remaining part of the proof is broken into three steps. First, we show that $(x^{\infty}(t), u^{\infty}(t))$ is a feasible solution to Problem B. Then, we prove the convergence of the cost function $\bar{J}^{N_i}(\bar{X}^*, \bar{U}^*)$ to the continuous cost function $J(x^{\infty}(\cdot), u^{\infty}(\cdot))$, and finally show that $(x^{\infty}(t), u^{\infty})$ is indeed an optimal solution of Problem B.

Step 1: To prove that $(x^{\infty}(t), u^{\infty}(t))$ is a feasible solution to Problem B, we first need to show that $(x^{\infty}(t), u^{\infty}(t))$ satisfies the state equation (1). By the contradiction argument, suppose $(x^{\infty}(t), u^{\infty}(t))$ is not a solution of the differential equation (1). Then there is a time $t' \in [-1, 1]$ so that

$$\dot{x}^{\infty}(t') - f(x^{\infty}(t'), u^{\infty}(t')) \quad \neq \quad 0$$

Since the LGL nodes t_k are dense with $N \to \infty$ [8], there exists a sequence k^{N_i} satisfying

$$0 < k^{N_i} < N_i \qquad \text{and} \qquad \lim_{i \to \infty} t_{k^{N_i}} = t'.$$

By assumption, $(x^{N_i}(t), \dot{x}^{N_i}(t), u^{N_i}(t))$ converge uniformly to $(x^{\infty}(t), \dot{x}^{\infty}(t), u^{\infty}(t))$; thus

$$\begin{aligned} & \dot{x}^{\infty}(t') - f(x^{\infty}(t'), u^{\infty}(t')) \\ = & \lim_{i \to \infty} (\dot{x}^{N_i}(t_{k^{N_i}}) - f(x^{N_i}(t_{k^{N_i}}), u^{N_i}(t_{k^{N_i}})) \neq 0 \end{aligned} \tag{20}$$

Because $x^N(t)$ is a N-th order polynomial, we have

$$\dot{x}^{N_i}(t_{k^{N_i}}) \; = \; \sum_{j=0}^{N_i} \bar{x}_j^* D_{k^{N_i} j}.$$

Thus from (11) and the fact that $(x^N(t), u^N(t))$ are the interpolating functions of $\{(\bar{x}_k^*, \bar{u}_k^*), 0 \leq k \leq N\}$, the following holds

$$\begin{aligned} & \lim_{i \to \infty} (\dot{x}^{N_i}(t_{k^{N_i}}) - f(x^{N_i}(t_{k^{N_i}}), u^{N_i}(t_{k^{N_i}})) \\ & = \lim_{i \to \infty} (N_i - 1)^{\frac{3}{2} - m} \; = \; 0 \end{aligned}$$

This contradicts (20); therefore, $(x^{\infty}(t), u^{\infty}(t))$ must be a solution of the differential equation (1).

The path constraint can be proved by the same contradiction argument. As for the end-point condition $e(x^{\infty}(-1), x^{\infty}(1)) = 0$, it follows directly from the convergence property, since

$$\begin{aligned} e(x^{\infty}(-1), x^{\infty}(1)) & = \lim_{i \to \infty} e(x^{N_i}(-1), x^{N_i}(1)) \\ & = \lim_{i \to \infty} e(\bar{x}_0^*, \bar{x}_{N_i}^*) \; = \; 0 \end{aligned}$$

Step 2: In this step, we will show that

$$\lim_{i \to \infty} \bar{J}^{N_i}(\bar{X}^*, \bar{U}^*) \; = \; J(x^{\infty}(\cdot), u^{\infty}(\cdot)), \tag{21}$$

where

$$\begin{aligned} \bar{J}^{N_i}(\bar{X}^*, \bar{U}^*) &= E(\bar{x}_0^*, \bar{x}_{N_i}^*) + \sum_{k=0}^{N_i} F(\bar{x}_k^*, \bar{u}_k^*) w_k \\ J(x^\infty(\cdot), u^\infty(\cdot)) &= E(x^\infty(-1), x^\infty(1)) + \\ &\int_{-1}^{1} F(x^\infty(t), u^\infty(t)) dt \end{aligned}$$

Since $(x^{N_i}(t), u^{N_i}(t))$ uniformly converges to $(x^\infty(t), u^\infty(t))$, we have,

$$\lim_{i\to\infty} |x^{N_i}(t_k) - x^\infty(t_k)| = \lim_{i\to\infty} |\bar{x}_k^* - x^\infty(t_k)| = 0 \quad (22)$$

$$\lim_{i\to\infty} |u^{N_i}(t_k) - u^\infty(t_k)| = \lim_{i\to\infty} |\bar{u}_k^* - u^\infty(t_k)| = 0 \quad (23)$$

uniformly in k. From this property, it is easy to conclude $(\bar{x}_k^*, \bar{u}_k^*)$ is bounded for all N_i and $0 \le k \le N_i$. Therefore, by the fact that $F(x, u)$ is continuously differentiable, there exists a constant $M > 0$ independent of N_i, such that

$$\begin{aligned} &|F(x^\infty(t_k), u^\infty(t_k)) - F(\bar{x}_k^*, \bar{u}_k^*)| \\ &\le M(|x^\infty(t_k) - \bar{x}_k^*| + |u^\infty(t_k) - \bar{u}_k^*|) \end{aligned}$$

for all $0 \le k \le N_i$. Furthermore, $F(x^\infty(t), u^\infty(t))$ is continuous in t. Thus, by Lemma 2, we have

$$\int_{-1}^{1} F(x^\infty(t), u^\infty(t)) dt = \lim_{i\to\infty} \sum_{k=0}^{N_i} F(x^\infty(t_k), u^\infty(t_k)) w_k$$

Therefore,

$$\begin{aligned} \int_{-1}^{1} F(x^\infty(t), u^\infty(t)) dt &= \lim_{i\to\infty} \left(\sum_{k=0}^{N_i} F(\bar{x}_k^*, \bar{u}_k^*) w_k + \right. \\ &\left. \sum_{k=0}^{N_i} \left[F(x^\infty(t_k), u^\infty(t_k)) - F(\bar{x}_k^*, \bar{u}_k^*) \right] w_k \right) \quad (24) \end{aligned}$$

From the uniform convergence of (22) and (23) and the property of w_k, $\sum_{k=0}^{N} w_k = 2$, we know that

$$\begin{aligned} &\lim_{i\to\infty} \left| \sum_{k=0}^{N_i} \left(F(x^\infty(t_k), u^\infty(t_k)) - F(\bar{x}_k^*, \bar{u}_k^*) \right) w_k \right| \\ &\le \lim_{i\to\infty} M \sum_{k=0}^{N_i} \left(|x^\infty(t_k) - \bar{x}_k^*| + |u^\infty(t_k)) - \bar{u}_k^*| \right) w_k = 0 \end{aligned}$$

Thus,

$$\int_{-1}^{1} F(x^\infty(t), u^\infty(t)) dt = \lim_{i\to\infty} \sum_{k=0}^{N_i} F(\bar{x}_k^*, \bar{u}_k^*) w_k \quad (25)$$

It is obvious that

$$\lim_{i\to\infty} E(\bar{x}_0^*, \bar{x}_{N_i}^*) = E(x^\infty(-1), x^\infty(1)) \quad (26)$$

Thus the limit in (21) follows from (25) and (26).

Step 3: Denote $(x^*(t), u^*(t))$ as any optimal solution of Problem B with the property that $x^*(t) \in W^{m,\infty}$, $m \ge 2$, (the optimal solution may not be unique). According to Theorem 1, there exists a sequence of feasible solutions, $(\tilde{x}_k^N, \tilde{u}_k^N)_{N=N_1}^\infty$, of Problem $\mathrm{B^N}$ that converge uniformly to $(x^*(t), u^*(t))$. Now, from (21) and the optimality of $(x^*(t), u^*(t))$ and $(\bar{x}_k^*, \bar{u}_k^*)$, we have

$$\begin{aligned} J(x^*(\cdot), u^*(\cdot)) &\le J(x^\infty(\cdot), u^\infty(\cdot)) \\ &= \lim_{i\to\infty} \bar{J}^{N_i}(\bar{X}^*, \bar{U}^*) \\ &\le \lim_{i\to\infty} \bar{J}^{N_i}(\tilde{X}, \tilde{U}). \quad (27) \end{aligned}$$

By using the same arguments as in Step 2, it is straightforward to show that

$$J(x^*(\cdot), u^*(\cdot)) = \lim_{i\to\infty} \bar{J}^{N_i}(\tilde{X}, \tilde{U}), \quad (28)$$

since $(\tilde{x}_k^N, \tilde{u}_k^N)_{N=N_1}^\infty$ converge uniformly to $(x^*(t), u^*(t))$. Equations (27) and (28) imply that

$$J(x^*(\cdot), u^*(\cdot)) = J(x^\infty(\cdot), u^\infty(\cdot))$$

This is equivalent to saying that $(x^\infty(\cdot), u^\infty(\cdot))$ is a feasible solution that achieves the optimal cost. Therefore, $t \mapsto (x^\infty(t), u^\infty(t))$ is an optimal solution to the optimal control Problem B. Thus, the conclusions in Theorem 2 follows. ■

Theorem 2 demonstrates that discrete Problem B^N is indeed a consistent approximation [20] to the continuous optimal control Problem B. In other words, *if the optimal solution of the discretized Problem B^N converges as N increases, then the limit point must be an optimal solution of the continuous Problem B*. Thus, under relatively mild conditions, Theorems 1-2 guarantee the existence and convergence of the discrete-time optimal solution to the continuous-time solution of the original problem.

V. Convergence of the Dual Variables

The convergence of dual variables is an extremely important issue in discrete approximations of optimal control problems as it provides insights on both the method of approximation and the resultant solution that would otherwise be unavailable from a consideration of the primal variables alone. For example, the discrepancy between the state and costate discretizations led Hager [11] to design new Runge-Kutta methods for control applications. Furthermore, in solving industrial-strength optimal control problems, verification and validation methods are crucial for safety, robustness and other issues. The Minimum Principle provides a plethora of such tests through integrals of motion (e.g. constancy of the lower Hamiltonian for autonomous systems) and other conditions. This is one reason why indirect methods (i.e. methods based on solving the necessary conditions arising from an application of the Minimum Principle) continue to be used; however, indirect methods are replete with many problems [2]. Thus, in designing direct methods that provide the appeal of indirect methods, a study of the convergence of dual variables takes center stage. In this section, we explore the link between the KKT multipliers and the discrete costates and clarify the covector mapping theorem of [24]. Throughout this section, we assume that Assumption 1 always holds.

A. Necessary Conditions for Problems B^N and B

Motivated by the results in [24], we use the discrete weights w_k to construct a Lagrangian for Problem B^N as

$$\begin{aligned} L^N &= \bar{J}^N + \sum_{k=0}^{N} \bar{\lambda}_k^T \Big(-\sum_{i=0}^{N} \bar{x}_i D_{ki} + f(\bar{x}_k, \bar{u}_k)\Big) w_k \\ &+ \bar{\nu}^T e(\bar{x}_0, \bar{x}_N) + \sum_{i=0}^{N} \bar{\mu}_k^T h(\bar{x}_k, \bar{u}_k) w_k \end{aligned}$$

where $\bar{\lambda}_k \in \mathbb{R}^{N_x}$, $\bar{\nu} \in \mathbb{R}^{N_e}$ and $\bar{\mu}_k \in \mathbb{R}^{N_h}$ are the KKT multipliers associated with Problem B^N. As a result of choosing a weighted inner product (1-form) for the construction of the Lagrangian, the KKT multipliers must be interpreted accordingly. Let, $\delta_P = (N-1)^{\frac{3}{2}-m_x}$. Then, a feasible point is called a KKT point if the KKT conditions are approximately satisfied,

$$\left|\left|\sum_{i=0}^{N} \bar{x}_i D_{ki} - f(\bar{x}_k, \bar{u}_k)\right|\right|_\infty \leq \delta_P, \tag{29}$$

$$h(\bar{x}_k, \bar{u}_k) \leq \delta_P \cdot \mathbf{1}, \quad ||e(\bar{x}_0, \bar{x}_N)||_\infty \leq \delta_P, \tag{30}$$

$$\left|\left|\frac{\partial L}{\partial \bar{u}_k}\right|\right|_\infty \leq \delta_D, \quad \left|\left|\frac{\partial L}{\partial \bar{x}_k}\right|\right|_\infty \leq \delta_D, \tag{31}$$

$$||\bar{\mu}_k \cdot h(\bar{x}_k, \bar{u}_k)||_\infty \leq \delta_D, \quad \bar{\mu}_k \geq -\delta_D \cdot \mathbf{1}, \tag{32}$$

where $k = 0, 1, \ldots, N$ and $\mathbf{1} = [1, \ldots, 1]^T$ with appropriate dimension and δ_D is a dual feasibility tolerance. A proper selection of δ_D will be apparent shortly. Part of the motivation for δ_D comes from the convergence criteria used in solving NLPs; see for example [9].

For the purpose of brevity, we omit a detailed derivation of an evaluation and subsequent simplification of (29) – (32); these steps can be found in [24]. The final result can be summarized as follows:

Problem $B^{N\lambda}$: Find a KKT point $(\bar{x}_k^*, \bar{u}_k^*, \bar{\lambda}_k^*, \bar{\mu}_k^*, \bar{\nu}^*)$, $k = 0, 1, \ldots, N$, of Problem B^{N} such that

$$\begin{aligned}
&\left|\left|\sum_{i=0}^{N} \bar{x}_i^* D_{ki} - f(\bar{x}_k^*, \bar{u}_k^*)\right|\right|_\infty \leq \delta_P \\
&\|e(\bar{x}_0^*, \bar{x}_N^*)\|_\infty \leq \delta_P, \qquad h(\bar{x}_k^*, \bar{u}_k^*) \leq \delta_P \cdot \mathbf{1} \\
&\left|\left|w_k \left[\sum_{i=0}^{N} \bar{\lambda}_i^* D_{ki} + F_x(\bar{x}_k^*, \bar{u}_k^*) + f_x^T(\bar{x}_k^*, \bar{u}_k^*)\bar{\lambda}_k^* + \right.\right.\right. \\
&\qquad \left.\left.\left. h_x^T(\bar{x}_k^*, \bar{u}_k^*)\bar{\mu}_k^*\right] + c_k\right|\right|_\infty \leq \delta_D \\
&\left|\left|w_k \left[F_u(\bar{x}_k^*, \bar{u}_k^*) + f_u^T(\bar{x}_k^*, \bar{u}_k^*)\bar{\lambda}_k^* + h_u^T(\bar{x}_k^*, \bar{u}_k^*)\bar{\mu}_k^*\right]\right|\right|_\infty \leq \delta_D \\
&||w_k\, \bar{\mu}_k^* \cdot h(\bar{x}_k^*, \bar{u}_k^*)||_\infty \leq \delta_D, \qquad \bar{\mu}_k^* \geq -\delta_D \cdot \mathbf{1},
\end{aligned}$$

where $c_i = 0$ for $2 \leq i \leq N-1$ and

$$\begin{aligned}
c_0 &= \bar{\lambda}_0^* + \frac{\partial E}{\partial x_0}(\bar{x}_0^*, \bar{x}_N^*) + (\frac{\partial e}{\partial x_0}(\bar{x}_0^*, \bar{x}_N^*))^T \bar{\nu}^* \\
c_N &= -\bar{\lambda}_N^* + \frac{\partial E}{\partial x_N}(\bar{x}_0^*, \bar{x}_N^*) + (\frac{\partial e}{\partial x_N}(\bar{x}_0^*, \bar{x}_N^*))^T \bar{\nu}^*.
\end{aligned}$$

The first-order necessary conditions for Problem B are based on Minimum Principle that uses the D-form of the Lagrangian of the Hamiltonian [12], $\overline{H}(x, u, \lambda, \mu) = H(x, u, \lambda) + \mu^T h(x, u)$, where $H(x, u, \lambda) = F(x, u) + \lambda^T f(x, u)$ is the control Hamiltonian, $\lambda(t)$ is the costate and $\mu(t)$ is the instantaneous KKT multiplier (covector) associated with the Hamiltonian Minimization Condition. Under suitable constraint qualifications [12], the necessary conditions for Problem B together with the state equation and constraints can be summarized as follows.

Problem B^λ: If $(x^*(t), u^*(t))$ is the optimal solution to Problem B^N, then there exist $(\lambda^*(t), \mu^*(t), v^*)$ such that

$$\begin{aligned}
\dot{x}^* &= f(x^*, u^*) \\
\dot{\lambda}^* &= -F_x(x^*, u^*) - f_x^T(x^*, u^*)\lambda^* - h_x^T(x^*, u^*)\mu^*(t) \\
0 &= F_u(x^*, u^*) + f_u^T(x^*, u^*)\lambda^* + h_u^T(x^*, u^*)\mu^*(t) \\
0 &= e(x^*(1), x^*(-1)) \\
0 &\geq h(x^*, u^*) \\
0 &= \mu^*(t)h(x^*(t), u^*(t)), \qquad \mu^*(t) \geq 0 \\
\lambda^*(-1) &= -E_{x(-1)}(x^*(-1), x^*(1)) - \\
&\quad e_{x(-1)}^T(x^*(-1), x^*(1))v^* \qquad (33) \\
\lambda^*(1) &= E_{x(1)}(x^*(-1), x^*(1)) + \\
&\quad e_{x(1)}^T(x^*(-1), x^*(1))v^* \qquad (34)
\end{aligned}$$

Remark 3: It is easy to observe that a PS discretization of Problem B^λ will not generate Problem $B^{N\lambda}$ although they appear to be similar. The discretization of Problem B^λ is denoted as Problem $B^{\lambda N}$ in Fig. 1 and illustrates that dualization and discretization are not necessarily commutative operations. As noted earlier, a similar observation has been made by Hager on Runge-Kutta methods.

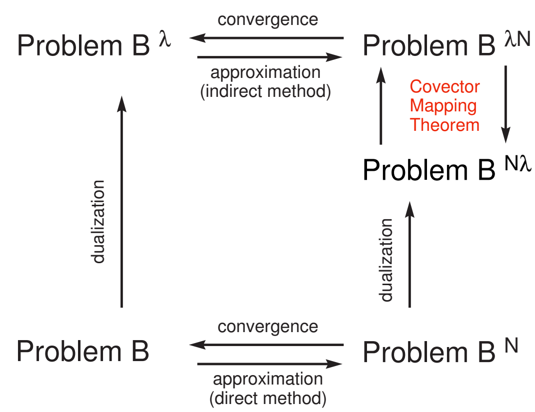

Fig. 1. Schematic for the covector mapping theorem [24].

The main points of Fig. 1 is illustrated by the following example. (which is a counter example to the widely-held notion that if the primals converge, the KKT multipliers associated with the discretized dynamic constraints converge to the costates).

Example 1: Minimize $J[x(\cdot), u(\cdot)] = x(2)$, subject to

$$\dot{x}(t) = u(t), \qquad t \in [0, 2] \tag{35}$$

$$x(0) = 0, \qquad u(t) \geq -1 \tag{36}$$

The dual feasibility conditions for (35)-(36) are

$$\begin{aligned}
\dot{\lambda}^*(t) &= 0, \qquad \lambda^*(2) = 1 \qquad (37) \\
\lambda^*(t) - \mu^*(t) &= 0 \\
\mu^*(t)(-u^*(t) - 1) &= 0, \qquad \mu^*(t) \geq 0
\end{aligned}$$

which uniquely determine the optimal solution as

$$x^*(t) = -t, \quad u^*(t) = -1, \quad \lambda^*(t) = \mu^*(t) = 1.$$

A numerical solution by the PS method with 20 nodes is shown in Figure 2. The left plot clearly shows that the primal variables $(\bar{x}_k^*, \bar{u}_k^*)$ coincide with the analytic solution $(x^*(t), u^*(t))$. On the other hand, the right plot shows that the KKT multipliers $\bar{\lambda}_k^*$, do not agree with the costate, $\lambda^*(t)$.

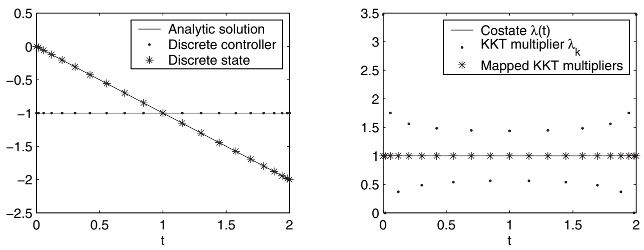

Fig. 2. Discrete solution by PS method with N=20.

These KKT multipliers were obtained in accordance with the weighted Lagrangian. If the standard inner-product is used, the disagreements between the multipliers are even worse [7]. Clearly, the convergence of the discretized primals does not imply the convergence of the KKT multipliers to the continuous costates. To clarify this point, consider the PS discretization of (35)-(36). For the purpose of clarity, we ignore the tolerances, δ_P and δ_D. This is further justified by the fact that the optimal continuous-time solutions being polynomials, the discretized problem can be posed exactly without introducing any infeasibility problem. Thus, an application of the method yields,

$$\text{Min.} \quad \bar{J}^N = \bar{x}_N, \qquad \text{subject to}$$

$$D\begin{pmatrix} \bar{x}_0 \\ \vdots \\ \bar{x}_N \end{pmatrix} = \begin{pmatrix} \bar{u}_0 \\ \vdots \\ \bar{u}_N \end{pmatrix} \tag{38}$$

$$\bar{x}_0 = 0 \tag{39}$$

$$\bar{u}_k \geq -1, \qquad 0 \leq k \leq N \tag{40}$$

It is easy to show that, for any N, the discretized problem admits a *unique globally optimal solution*: $\bar{u}_k^* = -1$, $\bar{x}_k^* = -t_k$, $0 \leq k \leq N$, where t_k are the LGL nodes. Thus, for any fixed N, the discrete optimal solution, $(\bar{x}_k^*, \bar{u}_k^*) = (-t_k, -1)$, can always be found and hence the convergence of the discrete solution to the continuous-time optimal solution is guaranteed. The left plot in Figure 2 demonstrates this point.

Next, the KKT conditions for the discrete problem are

$$D\begin{pmatrix} \bar{\lambda}_0^* \\ \bar{\lambda}_1^* \\ \vdots \\ \bar{\lambda}_{N-1}^* \\ \bar{\lambda}_N^* \end{pmatrix} = \begin{pmatrix} \bar{v} \\ 0 \\ \vdots \\ 0 \\ (\bar{\lambda}_N^* - 1)/w_N \end{pmatrix} \tag{41}$$

$$\bar{\lambda}_k^* = \bar{\mu}_k^*, \qquad 0 \leq k \leq N \tag{42}$$

$$\bar{\mu}_k^*(-u_k^* - 1) = 0, \qquad \bar{\mu}_k^* \geq 0 \tag{43}$$

where $\bar{v}$ is the multiplier associated with the initial condition. Since the constraint $\bar{u}_k \geq -1$ is always active at the optimal solution, $\bar{\mu}_k^*$ is undetermined in (43). In addition, it is easy to show that (41) has infinitely many solutions. In other words, the KKT multipliers are not unique although the optimal primal solution is unique. It is also straightforward to show that the linear independence constraint qualification is violated in this example but the weaker Mangasarian-Fromovitz constraint qualification [17] holds. Thus, the KKT multipliers exist but are not unique. This nonuniqueness persists even as $N \to \infty$. Uniqueness can be restored by supplying the missing condition, $\bar{\lambda}_N^* = 1$to the discrete dual feasibility conditions. This condition is obtained simply by comparing (41) with (37). With this additional condition, it is easy to see that the KKT conditions (41)—(43) admit a unique solution. This is plotted in the right plot of Figure 2 indicating a perfect match with the costate $\lambda(t)$.

B. The Augmented KKT Conditions

In the general case, comparing Problem $B^{\lambda N}$ with Problem $B^{N\lambda}$, it is apparent that the transversality conditions (33)-(34) are missing in the KKT conditions. Alternatively, the costate differential equations are not naturally collocated at the boundary points, -1 and 1. By restoring this information loss to the KKT conditions, the KKT multipliers can be mapped to the discretized covectors associated with Problem B^{λ}. More specially, the following conditions are needed in addition to the KKT conditions

$$\left\| -\bar{\lambda}_0^* - \frac{\partial E}{\partial x_0}(\bar{x}_0^*, \bar{x}_N^*) - (\frac{\partial e}{\partial x_0}(\bar{x}_0^*, \bar{x}_N^*))^T \bar{\nu}^* \right\|_\infty \leq \delta_D \tag{44}$$

$$\left\| \bar{\lambda}_N^* - \frac{\partial E}{\partial x_N}(\bar{x}_0^*, \bar{x}_N^*) - (\frac{\partial e}{\partial x_N}(\bar{x}_0^*, \bar{x}_N^*))^T \bar{\nu}^* \right\|_\infty \leq \delta_D \tag{45}$$

These equations generalize the "closure conditions" identified in [24]. They lead to a proof of Theorem 3 which clarifies the Covector Mapping Theorem [24].

Theorem 3 (Covector Mapping Theorem): Given any feasible solution, $t \mapsto (x, u, \lambda, \nu)$, for Problem B^{λ}, suppose $x(\cdot) \in W^{m_x,\infty}$ and $\lambda(\cdot) \in W^{m_\lambda,\infty}$ with $m_x, m_\lambda \geq 2$. Then, there exists a positive integer N_2 such that, for any $N > N_2$, the augmented KKT conditions, i.e., (29)-(30) plus (44)-(45), has a feasible solution with a primal feasibility tolerance of $\delta_P = (N-1)^{\frac{3}{2}-m_x}$ and a dual feasibility tolerance of $\delta_D = (N-1)^{\frac{3}{2}-m_\lambda}$.

The proof of this theorem is based on ideas similar to that of the proof of Theorem 1. For the purpose of brevity, we skip it.

Remark 4: In practice, we often observe the convergence of the primal variables, and as observed in Example 1, the KKT multipliers do not converge. In the absence of Theorem 3, the existence of a solution to the augmented KKT conditions was questionable. Theorem 3 guarantees the existence of solution to both the KKT conditions and the augmented KKT conditions. When multiple solutions exist for the KKT multipliers, the closure conditions, (44)-(45), act as a selection criterion in picking the proper set of KKT multipliers that constitute the subsequence which converges to the continuous-time covectors. In the event the KKT conditions admit a unique solution, the closure conditions do not introduce an infeasibility problem into the augmented KKT conditions.

We now establish a final theorem on the convergence of the sequence of the mapped dual variables. This is done in a manner similar to the analysis of the convergence of the primal variables. Let $(\bar{x}_k^*, \bar{u}_k^*, \bar{\lambda}_k^*, \bar{\mu}_k^*, \bar{\nu}^*)$, $k = 0, 1, \ldots, N$, be a solution to the augmented KKT conditions, i.e, Problem $B^{N\lambda}$ plus the closure conditions (44)-(45). Consider a sequence of the augmented KKT conditions with N increasing from N_2 to infinity. Correspondingly we get a sequence of discrete solutions $\{\bar{x}_k^*, \bar{u}_k^*, \bar{\lambda}_k^*, \bar{\mu}_k^*, \bar{\nu}^*\}_{N=N_2}^{\infty}$. Furthermore, denote $(x^N(t), \lambda^N(t))$ as the N-th order interpolating polynomials of $(\bar{x}_k^*, \bar{\lambda}_k^*)$, and $(u^N(t), \mu^N(t))$ as any interpolating function of $(\bar{u}_k^*, \bar{\mu}_k^*)$, i.e.

$$\begin{aligned} x^N(t) &= \sum_{k=0}^N \bar{x}_k^* \phi_k(t), \qquad u^N(t) = \sum_{k=0}^N \bar{u}_k^* \psi_k(t), \\ \lambda^N(t) &= \sum_{k=0}^N \bar{\lambda}_k^* \phi_k(t), \qquad \mu^N(t) = \sum_{k=0}^N \bar{\mu}_k^* \psi_k(t), \end{aligned}$$

where $\phi_k(t)$ is the Lagrange interpolating polynomial defined by (4) and $\psi_k(t)$ is any continuous function such that $\psi_k(t_j) = 1$, if $k = j$ and $\psi_k(t_j) = 0$, if $k \neq j$. For instance, $\psi(t)$ can be a linear or spline interpolant.

Assumption 2: Let the primal and dual feasibility tolerances be chosen as in Theorem 3. Suppose the sequences $\{x^N(t_0)\}_{N=N_2}^\infty$, $\{\lambda^N(t_0)\}_{N=N_2}^\infty$ and $\{\bar{\nu}^*\}_{N=N_2}^\infty$ converge as $N \to \infty$; denote their limits as $(x_0^\infty, \lambda_0^\infty, \bar{\nu}^\infty)$.

Theorem 4: Let $\{x^N(t), u^N(t), \lambda^N(t), \mu^N(t)\}_{N=N_2}^\infty$ be a sequence of interpolating functions constructed from optimal solutions, $(\bar{x}_k^*, \bar{u}_k^*, \bar{\lambda}_k^*, \bar{\mu}_k^*)$, $k = 0, 1, \ldots, N$, to Problem B^{N}. Suppose Assumption 2 holds. Let continuous functions $(\eta(t), u^\infty(t), \rho(t), \mu^\infty(t))$ be any uniform accumulation point of the sequence $(\dot{x}^N(t), u^N(t), \dot{\lambda}^N(t), \mu^N(t))$. Then the functions $(x^\infty(t), u^\infty(t), \lambda^\infty(t), \mu^\infty(t))$ must satisfy all the necessary conditions indicated by Problem B^λ, where

$$\begin{aligned} x^\infty(t) &= \int_{-1}^t \eta(\tau) d\tau + x_0^\infty \\ \lambda^\infty(t) &= \int_{-1}^t \rho(\tau) d\tau + \lambda_0^\infty. \end{aligned}$$

This theorem can be proved in the same manner as the proof of Theorem 2 and is therefore omitted.

Remark 5: Theorem 4 completes all the associations identified in Fig. 1. Thus, although a direct PS method is used to solve the optimal control problem, the results of Theorem 3-4 indicate that there is essentially no distinction between direct and indirect PS methods. Even more appealing is the fact that the well-known ease and robustness of direct methods can now be used to solve problems while still maintaining a direct link to the Minimum Principle but without all the difficulties associated with solving problems by an indirect method.

Remark 6: As indicated earlier, a Legendre PS method is available through the software package, DIDO [22]. DIDO has been publicly available since about 2001. Recently, the Legendre PS method became available as part of the NASA-developed software package, OTIS [18] (Version 4.0). Unlike DIDO, OTIS has a substantial number of additional tools for generating quick solutions to aerospace trajectory optimization problems.

REFERENCES

[1] J. T. Betts, N. Biehn and S. L. Campbell, Convergence of nonconvergent IRK discretizations of optimal control problems with state inequality constraints, *SIAM J. Sci. Comp.*, Vol. 23, No. 6, pp. 1981-2007, 2002.

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

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

[4] A. L. Dontchev and W. W. Hager, The Euler approximation in state constrained optimal control, *Math. Comp.,* Vol. 70, pp. 173-203, 2000.

[5] G. Elnagar, M. A. Kazemi and M. Razzaghi, The pseudospectral Legendre method for discretizing optimal control problems, *IEEE Trans. Auto. Cont.* Vol. 40, pp. 1793-1796, 1995.

[6] F. Fahroo and I. M. Ross, Costate estimation by a Legendre pseudospectral method, *J. of Guid., Cont. & Dyn.*, Vol. 24, No. 2, pp. 270-277, 2001.

[7] F. Fahroo and I. M. Ross, On discrete-time optimality conditions for pseudospectral methods, *Proc. AIAA/AAS Astro. Conf.*, Keystone, CO, August 2006. AIAA-2006-6304.

[8] G. Freud, *Orthogonal polynomials,* Pergamon Press, 1971.

[9] P. E. Gill, W. Murray and M. A. Saunders, SNOPT: an SQP algorithm for large-scale constrained optimization, *SIAM Review*, Vol. 47, No. 1, 2005, pp. 99-131.

[10] Q. Gong, W. Kang and I. M. Ross, A pseudospectral method for the optimal control of constrained feedback linearizable systems, *IEEE Trans. Auto. Cont.*, Vol. 51, No. 7, July 2006, pp. 1115-1129.

[11] W. W. Hager, Runge-Kutta methods in optimal control and the transformed adjoint system, *Numerische Mathematik,* Vol. 87, pp. 247-282, 2000.

[12] R. F. Hartl, S. P. Sethi and R. G. Vickson, A survey of the maximum principles for optimal control problems with state constraints, *SIAM Rev.,* Vol. 37, No. 2, pp. 181-218, 1995.

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

[14] S. I. Infeld and W. Murray, Optimization of stationkeeping for a Libration point mission, *Proc. AAS Spaceflight Mech. Mtg.*, Maui, HI, February 2004. AAS 04-150.

[15] P. Lu, H. Sun and B. Tsai, Closed-Loop endoatmospheric ascent guidance, *J. Guid., Cont. & Dyn.*, Vol. 26, No. 2, pp. 283-294, 2003.

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

[17] J. Nocedal and S. J. Wright, *Numerical Optimization*, Springer-Verlag, New York, 1999.

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

[19] J. Pietz and N. Bedrossian, Moemtum dumping using only CMGs, *Proc. AIAA GNC Conf.*, Austin, TX, 2003.

[20] E. Polak, *Optimization: Algorithms and Consistent Approximations*, Springer-Verlag, Heidelberg, 1997.

[21] J. Rea, Launch vehicle trajectory optimization using a Legendre pseudospectral method, *Proc. AIAA GNC Conf.*, Austin, TX, Paper No. AIAA 2003-5640, August 2003.

[22] I. M. Ross, User's manual for DIDO: A MATLAB application package for solving optimal control problems, Technical Report 04-01.0, Tomlab Optimization Inc, February 2004.

[23] I. M. Ross and F. Fahroo, Pseudospectral methods for optimal motion planning of differentially flat systems, *IEEE Trans. Auto. Cont.*, Vol. 49, No. 8, pp. 1410-1413, August 2004.

[24] I. M. Ross and F. Fahroo, Legendre pseudospectral approximations of optimal control problems, *Lec. Notes in Cont. & Info. Sc.*, Eds. W. Kang et al, Vol. 295, Springer-Verlag, pp. 327-342, New York, 2003.

[25] I. M. Ross and F. Fahroo, A unified framework for real-time optimal control, *Proc. IEEE CDC*, Maui, December, pp. 2210-2215, 2003.

[26] I. M. Ross, P. Sekhavat, A. Fleming and Q. Gong, Pseudospectral feedback control: Foundations, examples and experimental results, *Proc. AIAA GNC Conf.*, 21-24 August 2006, Keystone, CO.

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

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

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

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