Title
A pseudospectral method for the optimal control of constrained feedback linearizable systems

Journal
IEEE Transactions on Automatic Control, 51(7)

ISSN
0018-9286

Authors
Gong, Qi
Kang, Wei
Ross, IM

Publication Date
2006-07-01

DOI
10.1109/tac.2006.878570

Peer reviewed

A Pseudospectral Method for the Optimal Control of Constrained Feedback Linearizable Systems

Qi Gong, *Member, IEEE*, Wei Kang, *Member, IEEE*, and I. Michael Ross

***Abstract*—We consider the optimal control of feedback linearizable dynamical systems subject to mixed state and control constraints. In general, a linearizing feedback control does not minimize the cost function. Such problems arise frequently in astronautical applications where stringent performance requirements demand optimality over feedback linearizing controls. In this paper, we consider a pseudospectral (PS) method to compute optimal controls. We prove that a sequence of solutions to the PS-discretized constrained problem converges to the optimal solution of the continuous-time optimal control problem under mild and numerically verifiable conditions. The spectral coefficients of the state trajectories provide a practical method to verify the convergence of the computed solution. The proposed ideas are illustrated by several numerical examples.**

***Index Terms*—Constrained optimal control, pseudospectral, nonlinear systems.**

I. INTRODUCTION

IT IS well known [7], [8], [47] that it is extremely difficult to analytically solve a state- and control-constrained nonlinear optimal control problem. The main difficulty arises 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 past decades, many computational methods have been developed for solving nonlinear optimal control problems. For instance, in [8], various numerical methods such as neighboring extremal methods, gradient methods and quasilinearization methods are discussed in detail. An update on these methods, along with extensive numerical results is presented in [7]. In [28], a generalized gradient method is proposed for constrained optimal control problems. In [11], the feasibility and convergence of a modified Euler discretization method is proved. A unified approach based on a piecewise constant approximation of the control is proposed in [23].

Numerical methods for solving nonlinear optimal control problems are typically described under two categories: direct methods and indirect methods [3]. Historically, many early numerical methods were based on finding solutions to satisfy a set of necessary optimality conditions resulting from Pontryagin's Maximum Principle [3][35]. These methods are collectively called indirect methods. There are many successful implementations of indirect methods including launch vehicle trajectory design, low-thrust orbit transfer, etc. [3], [6], [9]. Although indirect methods enjoy some nice properties, they also suffer from many drawbacks [2]. For instance, the boundary value problem resulting from the necessary conditions are extremely sensitive to initial guesses [8][2]. In addition, these necessary conditions must be explicitly derived—a labor-intensive process for complicated problems that requires an in-depth knowledge of optimal control theory.

Over the last decade, an alternative approach based on discrete approximations has gained wide popularity [27], [2], [13], [15], [26], [46] as a result of significant progress in large-scale computation and robustness of the approach. The essential idea of this method is to discretize the optimal control problem and solve the resulting large-scale finite-dimensional optimization problem. These types of methods are known as direct methods whose roots can be traced back to the works of Bernoulli and Euler [39]. The simplicity of direct methods belies a wide range of deeply theoretical issues 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], [30], [33], [37].

Considering the widespread use of discrete approximations, it might appear to a novice that theoretical questions regarding the existence of a solution and convergence of the approximations have been answered satisfactorily. While this is somewhat true of Eulerian methods [29], [12], [13], [35], [11] corresponding results for higher-order methods are not only absent, but a number of interesting results are reported in the literature. For example, Hager [26] has shown that a "convergent" Runge–Kutta method does not converge to the continuous optimal solution despite the fact that it satisfies the standard conditions in the Butcher tableau. On the other hand, Betts *et al.* [1] show that a nonconvergent Runge–Kutta method converges for optimal control problems. Thus, it is not surprising that even for Eulerian methods, significant restrictions and assumptions are necessary for proofs of convergence, particularly for state-constrained problems [13]. Note that these issues are quite different from those that were raised in the early days of optimal control as summarized in [35]. While Eulerian methods are widely studied and useful for a theoretical understanding of discrete approximations, they are not practical for solving industrial strength problems, especially the so-called multiagent problems that require real-time solutions

Manuscript received July 6, 2005; revised November 30, 2005. Recommended by Associate Editor A. Astolfi. This work was supported by the Secretary of the Air Force, the Naval Postgraduate School, and the Jet Propulsion Laboratory. The work of Q. Gong was supported by the National Academies' National Research Council postdoctoral program.

Q. Gong and I. M. Ross are with the Department of Mechanical and Astronautical Engineering, Naval Postgraduate School, Monterey, CA, 93943 (e-mail: qgong@nps.edu; imross@nps.edu).

W. Kang is with the Department of Applied Mathematics, Naval Postgraduate School, Monterey, CA, 93943 (e-mail: wkang@nps.edu).

Digital Object Identifier 10.1109/TAC.2006.878570

TABLE I
COMPARISON OF DIFFERENT DISCRETIZATION METHODS

Pseudospectral			Hermite-Simpson			Euler		
Nodes	Error	Time	Nodes	Error	Time	Nodes	Error	Time
10	1.5458×10^{-3}	0.155s	10	1.0747×10^{-2}	0.156s	100	1.5071×10^{-2}	0.782s
12	1.2366×10^{-4}	0.264s	25	1.6482×10^{-3}	0.250s	200	7.5681×10^{-3}	3.117s
14	8.6362×10^{-6}	0.263s	40	6.3733×10^{-4}	0.452s	300	5.0534×10^{-3}	8.182s
16	5.4438×10^{-7}	0.279s	55	3.3551×10^{-4}	0.779s	400	3.7915×10^{-3}	23.037s
18	3.2477×10^{-8}	0.326s	70	2.0740×10^{-4}	1.465s	500	3.0351×10^{-3}	37.451s

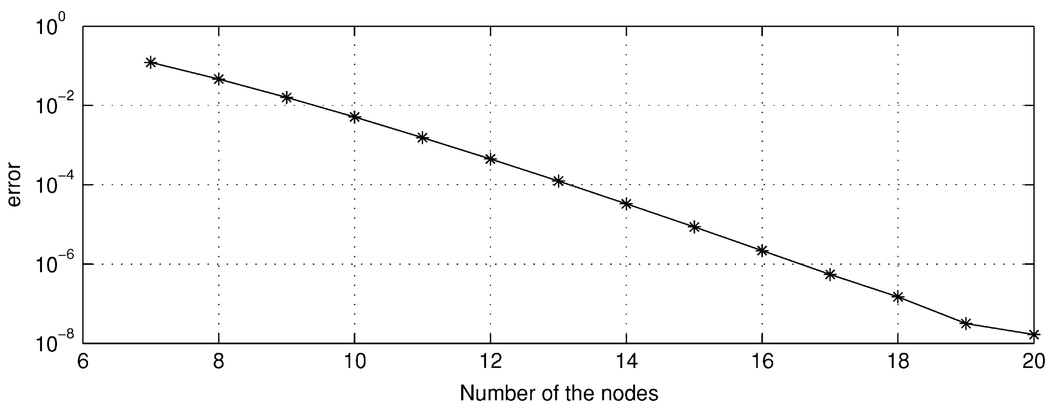


Fig. 1. Demonstrating the exponential convergence rate of a PS method.

[43]. Among many, one of the reasons for their limitation is that they generate a much larger-scale optimization problem than say, a higher order scheme like a Runge–Kutta method [14].

In this paper we focus on direct pseudospectral (PS) methods. PS methods were largely developed in the 1970s for solving partial differential equations arising in fluid dynamics and meteorology [10], and quickly became "one of the big three technologies for the numerical solution of PDEs" [45]. During the 1990s, PS methods were introduced for solving optimal control problems [15]–[18]; and since then, have gained considerable attention [19], [30], [33], [37], [44], [48]. One of the main reasons for the popularity of PS methods is that they demonstrably offer an exponential convergence rate for the approximation of analytic functions [45] while providing Eulerian-like simplicity. Thus, for a given error bound, PS methods generate a significantly smaller scale optimization problem when compared to other methods on problems with highly smooth solutions. This property is particularly attractive for control applications as it places real-time computation within easy reach of modern computational power [43]. To illustrate these points, consider the nonlinear optimal control problem shown in the equation at the bottom of the page.

Example 1: It is straightforward to show that the exact optimal control is given by $u^*(t) = -(8/(2+t)^3)$. Now, consider three different discretizations of the problem: Euler, Hermite–Simpson [27] and PS. Part of the reason for choosing the Hermite–Simpson discretization scheme for comparison is because it is a widely used method and forms the basis of well-known commercial optimal control software packages such as OTIS [34] and SOCS [2]. Since different implementations, in terms of software and hardware, of the same discretization method may cause variations in the results, the simulations reported in Table I are to be considered as illustrative. The column labeled "Nodes" denotes the number of nodes used in the discretization; the "Error" column denotes the maximum error in control between the discrete and exact solutions; and, the "Time" column represents the computer run-time for the calculation of the solution. All of the discretization methods are implemented on the same computational hardware (Pentium 4, 2.4-GHz with 256 MB of RAM) and software (MATLAB 6.5 under Windows XP). The resulting nonlinear programming problems were solved by the sequential quadratic programming (SQP) method of SNOPT [22]. A quick glance at the table illustrates that the PS method achieves a much higher

$$\begin{cases} \text{Minimize} & J[x(\cdot), u(\cdot)] = 4x_1(2) + x_2(2) + 4\int_0^2 u^2(t)\,dt \\ \text{Subject to} & \dot{x}_1(t) = x_2^3(t) \\ & \dot{x}_2(t) = u(t) \\ & (x_1(0), x_2(0)) = (0, 1) \end{cases}$$

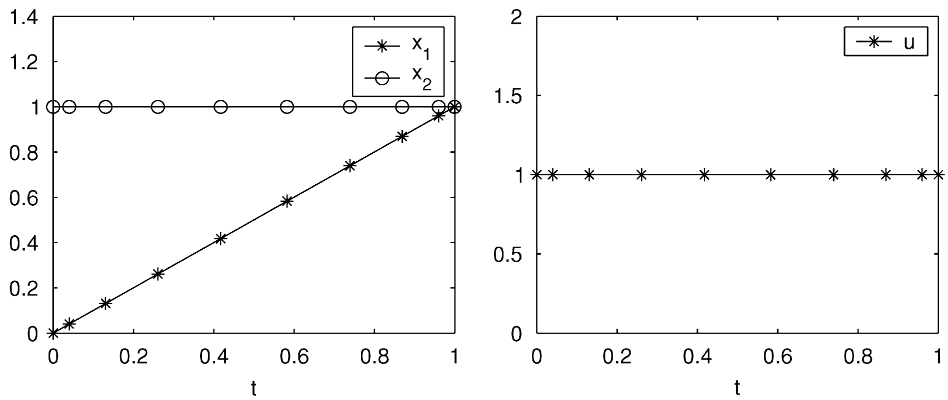


Fig. 2. Discrete optimal solution for Example 2 by a PS method with 10 nodes.

accuracy with a significantly smaller number of nodes and run time than the other methods. The third column of Table I demonstrates the linear convergence rate of the Euler method as has been proved in [13]. For the Hermite–Simpson method, our numerical experiment shows that the convergence rate is $N^{-1.97}$. From the logarithmically scaled ordinate of Fig. 1, it is not difficult to see the exponential convergence rate of the PS method.

This simple example illustrates the spectral convergence property of PS methods and its viability for real-time computation [43]. Experiments with other problems show similar behavior; see, for example, Lu *et al.* [33] who report a similar observation with regards to a launch vehicle ascent problem which is significantly more complicated than Example 1. Computational experience suggests that PS methods may also have better robustness than other discretizations. This is quickly illustrated by the following example adapted from [46].

Example 2:

$$\begin{cases} \text{Minimize} & J[x(\cdot),u(\cdot)] = \int_0^1 x_2 u \, dt \\ \text{Subject to} & \dot{x}_1(t) = x_2(t) \\ & \dot{x}_2(t) = -x_2(t) + u(t) \\ & x_2(t) \geq 0 \\ & 0 \leq u \leq 2 \\ & (x_1(0), x_2(0)) = (0,1) \\ & (x_1(1), x_2(1)) = (1,1). \end{cases} \tag{I.1}$$

This problem describes a particle moving under friction, where x_1 is the position, x_2 is the velocity and u is the applied force. The optimal control problem is to minimize the total amount of work done. From physical considerations or a direct application of the Minimum Principle, it is easy to verify that the optimal control is a constant that is equal to the amount of force required to maintain the initial speed. Since the optimal solution is smooth, a PS method is expected to achieve superior performance; this is demonstrated in Figs. 2 and 3 shows the results for the Hermite–Simpson and Euler discretizations. The nonlinear programming (NLP) solver and the computer used are exactly the same as those used in the PS method. It is clear that the discrete optimal controllers resulting from the Hermite–Simpson and the Euler methods are incorrect. One might expect an improvement in the solution with an increase in the number of nodes. That this is not true is also demonstrated in Fig. 3(b) and (d), where additional nodes are added. This phenomenon was reported in [46] with similar conclusions for the midpoint and trapezoidal discretizations. Furthermore, the phenomenon was also invariant with respect to two different types of NLP solvers: SNOPT [22] (a sequential quadratic programming solver) and LOQO [46] (an interior point solver). The failure of traditional methods on this problem is due to the well-known problems of computation along singular arcs [2].

As a result of similar experiences with many other problems at NASA, the next generation of the OTIS software package [34] will have 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 software package, DIDO [38] (developed by one of the authors), uses PS methods exclusively for solving optimal control problems. Part of the appeal of PS methods for control application is that it offers a ready approach to exploiting differential-geometric properties of a control system such as convexity and differential flatness [40]. In this paper, we exploit the normal form of feedback linearizable dynamics.

Much of the prior work on PS methods for control has largely focused on the development of the techniques, algorithms, and engineering applications. Rigorous convergence proofs and error estimation formulas are essentially unavailable. This is, in part, because PS methods for control are of recent vintage when compared to, say, the Runge–Kutta methods. In addition, standard convergence theorems frequently employed in the numerical analysis of differential equations are not applicable to discretizations of optimal control problems as has been noted by Betts *et al.* [1] and Hager [26]. Hager [26] has derived additional conditions for convergence of Runge–Kutta methods by eliminating the discrepancies that arise in the discrete costate equations. On the other hand, Betts *et al.* [1] show that "nonconvergent" Runge–Kutta (implicit) methods converge for discretizations of optimal control problems. What has emerged in recent years is that the convergence theorems of Eulerian methods are not portable to higher order methods. Furthermore, with regards to PS methods, its marked differences with other methods imply that standard convergence theorems are not applicable for PS discretizations. Thus, a new approach is needed to address some fundamental questions. In this paper,

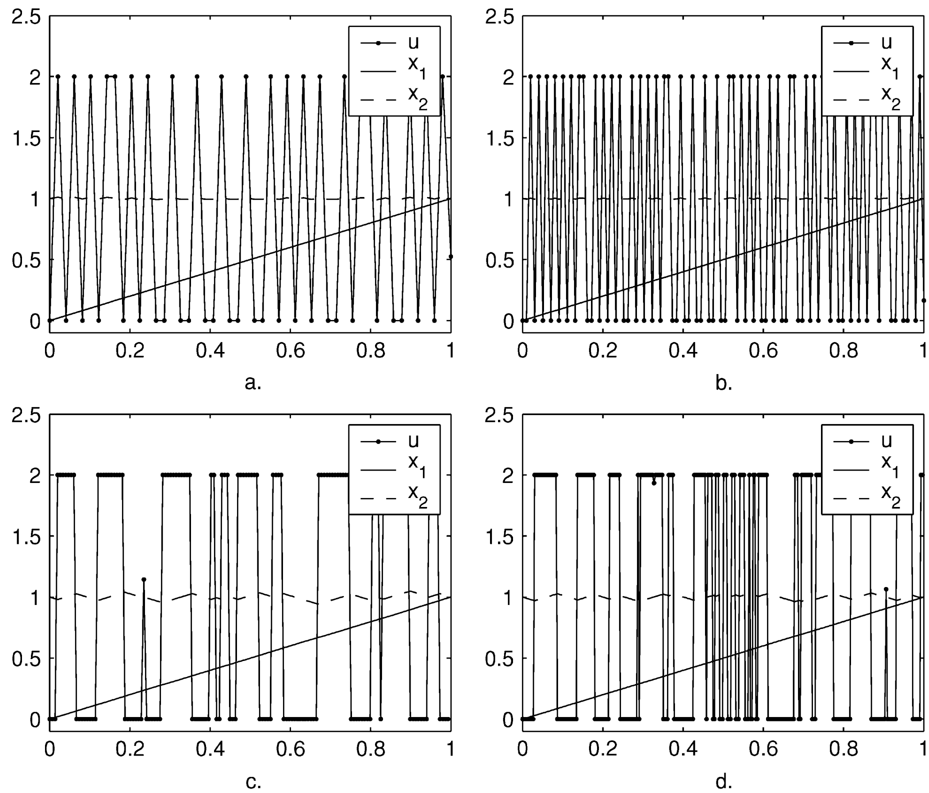


Fig. 3. Discrete optimal solutions for Example 2 by Hermite-Simpson and Euler methods.

we address some of these basic questions. For example, does the discretized problem always have a solution if a solution to the continuous-time problem 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 [43].

If a dynamical system can be written in normal form, it permits a modification of the standard pseudospectral method [15], [42] in a manner that is similar to dynamic inversion. That is, we seek polynomial approximations of the state trajectories while the controls are determined by an exact satisfaction of dynamics. This modification of a pseudospectral method permits us to prove sufficient conditions for the existence and convergence of PS discretizations. Furthermore, our method allows one to easily incorporate state and control constraints including mixed state and control constraints. Note that we do not linearize the dynamics by feedback control; rather, we find the optimal control for a generic cost function. Such problems are particularly common in astronautical applications where stringent performance requirements demand that the control be optimal rather than merely feasible as implied by the linearizing control. We show that, under mild and verifiable conditions, the PS discretized optimization problem always has a feasible solution. This is in sharp contrast to a noncontrol-affine dynamics which requires a relaxation of the dynamical constraints [24], [25]. Further, we show that the numerical solution converges to the solution of the original continuous-time constrained optimal control problem. We illustrate our methods using several examples including one from robotics for which no closed-form solution is presented in the literature.

The paper is organized as follows. In Section II, we briefly present a PS discretization method for constrained nonlinear optimal control problems. Section III contains the main results regarding existence and convergence of the PS discretization. In Section IV, some remarks and extensions of the main results are presented. Finally, in Section V, we illustrate a few key points by computing solutions to some specific problems.

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

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

A definition of distributional derivatives can be found in the Appendix of [10]. For notational ease, we suppress the dependence of $W^{m,p}$ on vector-valued functions.

II. The Problem and Its Discretization

We consider the following mixed, state- and control constrained nonlinear Bolza problem (Problem B) with single input feedback linearizable dynamics.

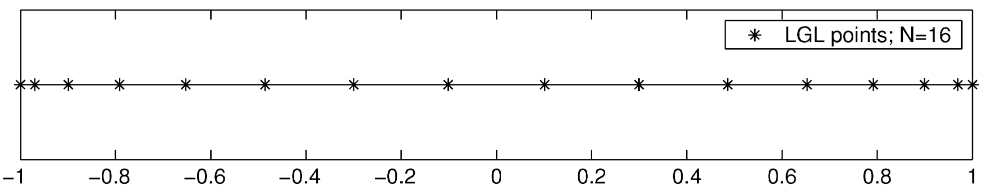


Fig. 4. Distribution of LGL nodes.

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

$$J[x(\cdot), u(\cdot)] = \int_{-1}^{1} F(x(t), u(t))\, dt + E(x(-1), x(1)) \tag{II.2}$$

subject to the dynamics

$$\begin{aligned} \dot{x}_1(t) &= x_2(t) \\ &\vdots \\ \dot{x}_{r-1}(t) &= x_r(t) \\ \dot{x}_r(t) &= f(x(t)) + g(x(t))u(t) \end{aligned} \tag{II.3}$$

endpoint conditions

$$e(x(-1), x(1)) = 0 \tag{II.4}$$

and path constraints

$$h(x(t), u(t)) \leq 0 \tag{II.5}$$

where $x \in \mathbb{R}^r, u \in \mathbb{R}$, and $F : \mathbb{R}^r \times \mathbb{R} \to \mathbb{R}$, $E : \mathbb{R}^r \times \mathbb{R}^r \to \mathbb{R}, f : \mathbb{R}^r \to \mathbb{R}, g : \mathbb{R}^r \to \mathbb{R}e : \mathbb{R}^r \times \mathbb{R}^r \to \mathbb{R}^{N_e}$ and $h : \mathbb{R}^r \times \mathbb{R} \to \mathbb{R}^s$ are Lipschitz continuous (over the domain) with respect to their arguments. For controllability reasons, we assume $g(x) \neq 0$ for all x. In addition to these standard assumptions, we assume that an optimal solution $(x^*(\cdot), u^*(\cdot))$ exists with the optimal state, $x_r^*(\cdot) \in W^{m,\infty},\ m \geq 2$. Note that, if $x_r^*(t)$ is C^1 and $\dot{x}_r^*(t)$ has bounded derivative everywhere except for finitely many points on the closed interval $t \in [-1, 1]$, then $x_r^*(\cdot)$ belongs to $W^{2,\infty}$. On the other hand, by Sobolev's Imbedding Theorems [10], any function $x_r^*(\cdot) \in W^{m,\infty}, m \geq 2$ must have continuous $(m-1)$th order classical derivatives on $[-1, 1]$. Therefore, this condition requires the optimal state $x_r^*(t)$ be at least continuously differentiable.

Remark 1: Pseudospectral methods are not limited to dynamical systems in normal form; in fact, they are applicable to far more general nonlinear systems; see for example, [42], [24], and the references contained therein. What the normal form facilitates is a proof of convergence of the computed system trajectory without dualizing the problem.

Remark 2: In Problem B, we assume the time interval to be fixed at $[-1, 1]$ in order to facilitate a simpler bookkeeping in using the Legendre pseudospectral method. If the physical time domain of the problem is not $[-1, 1]$, it can always be projected to the computational domain $[-1, 1]$ by a simple linear transformation [15], [20].

In the Legendre pseudospectral approximation of Problem B, the basic idea is to approximate $x(t)$ by an N-th order Lagrange interpolating polynomial, $x^N(t)$, based on the Legendre–Gauss–Lobatto (LGL) quadrature nodes. Let $t_0 = -1 < t_1 < \cdots < t_N = 1$ be the LGL nodes defined as

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

where $\dot{L}_N(t)$ is the derivative of the Nth-order Legendre polynomial $L_N(t)$. The distribution of the LGL nodes is illustrated in Fig. 4. Note that the node distribution is not uniform. The high density of nodes near the end points is one of the key properties of PS discretizations in that it effectively prevents the Runge phenomenon. Computational advantages of such nonuniformly distributed quadrature nodes can be found in [5], [10], and [45].

Let $\bar{x}_k^N$ and $\bar{u}_k^N$ be an approximation of a feasible solution $(x(t), u(t))$ evaluated at the node t_k. Then, $x^N(t)$ is used to approximate $x(t)$ by

$$x(t) \approx x^N(t) = \sum_{k=0}^{N} \bar{x}_k^N \phi_k(t) \tag{II.6}$$

where $\phi_k(t)$ is the Lagrange interpolating polynomial defined by

$$\phi_k(t) = \frac{1}{N(N+1)L_N(t_k)} \frac{(t^2-1)\dot{L}_N(t)}{t - t_k}. \tag{II.7}$$

It is readily verifiable that $\phi_k(t_j) = 1$, if $k = j$ and $\phi_k(t_j) = 0$, if $k \neq j$. The precise nature of the approximation indicated in (II.6) is the main focus of this paper. From (II.3), the control that generates the approximate state is given by

$$u^N(t) = \frac{\dot{x}_r^N(t) - f(x^N(t))}{g(x^N(t)).} \tag{II.8}$$

Note that $u^N(t)$ is not necessarily a polynomial and hence differs from a standard pseudospectral approximation. The derivative of $x_i^N(t)$ at the LGL node t_k is given by

$$\dot{x}_i^N(t_k) = \sum_{j=0}^{N} D_{kj}\, x_i^N(t_j), \qquad i = 1, 2, \ldots, r \tag{II.9}$$

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

$$D_{ik} = \begin{cases} \frac{L_N(t_i)}{L_N(t_k)}\frac{1}{t_i-t_k}, & \text{if } i \neq k \\ -\frac{N(N+1)}{4}, & \text{if } i=k=0 \\ \frac{N(N+1)}{4}, & \text{if } i=k=N \\ 0, & \text{otherwise.} \end{cases} \tag{II.10}$$

Throughout this paper, we use the "bar" notation to denote corresponding variables in the discrete space, and the superscript N to denote the number of nodes used in the discretization. Thus, let

$$\bar{x}_0^N = \begin{pmatrix} \bar{x}_{10}^N \\ \vdots \\ \bar{x}_{r0}^N \end{pmatrix}, \bar{x}_1^N = \begin{pmatrix} \bar{x}_{11}^N \\ \vdots \\ \bar{x}_{r1}^N \end{pmatrix}, \cdots, \bar{x}_N^N = \begin{pmatrix} \bar{x}_{1N}^N \\ \vdots \\ \bar{x}_{rN}^N \end{pmatrix}.$$

Note that the subscript in $\bar{x}_k^N \in \mathbb{R}^{N_r}$ denotes an evaluation of the approximate state, $x^N(t) \in \mathbb{R}^{N_r}$, at the node t_k. whereas $x_k(t)$ denotes the kth component of the exact state.

With these preliminaries, it is apparent that the approximate solutions must satisfy the following nonlinear algebraic equations:

$$D\begin{pmatrix} \bar{x}_{10}^N \\ \vdots \\ \bar{x}_{1N}^N \end{pmatrix} = \begin{pmatrix} \bar{x}_{20}^N \\ \vdots \\ \bar{x}_{2N}^N \end{pmatrix}$$
$$\vdots$$
$$D\begin{pmatrix} \bar{x}_{r-1,0}^N \\ \vdots \\ \bar{x}_{r-1,N}^N \end{pmatrix} = \begin{pmatrix} \bar{x}_{r0}^N \\ \vdots \\ \bar{x}_{rN}^N \end{pmatrix}$$
$$D\begin{pmatrix} \bar{x}_{r0}^N \\ \vdots \\ \bar{x}_{rN}^N \end{pmatrix} = \begin{pmatrix} f\left(\bar{x}_0^N\right) + g\left(\bar{x}_0^N\right)\bar{u}_0^N \\ \vdots \\ f\left(\bar{x}_N^N\right) + g\left(\bar{x}_N^N\right)\bar{u}_N^N \end{pmatrix} \tag{II.11}$$

for feasibility with respect to the dynamics. In a standard pseudospectral method, it is quite common [43], [20], [15], [40] to discretize the mixed state- and control constraints as

$$h\left(\bar{x}_k^N, \bar{u}_k^N\right) \leq 0, \qquad k = 0, 1, \ldots, N. \tag{II.12}$$

Here, to guarantee feasibility we propose the following relaxation:

$$h\left(\bar{x}_k^N, \bar{u}_k^N\right) \leq (N-r)^{-m+\frac{3}{2}} \cdot \mathbf{1}, \qquad k = 0, 1, \ldots, N \tag{II.13}$$

where $\mathbf{1}$ denotes $[1, \ldots, 1]^T$. When N tends to infinity, the difference between conditions (II.12) and (II.13) vanishes, since by assumption, $m \geq 2$. The purpose of this relaxation will be evident in Section III. Similarly, we relax the endpoint condition $e(x(-1), x(1)) = 0$, to an inequality, i.e.,

$$\left\|e(\bar{x}_0^N, \bar{x}_N^N)\right\|_\infty \leq (N-r)^{-m+\frac{3}{2}}. \tag{II.14}$$

Remark 3: The right-hand side of (II.13) and (II.14) can be set to $(N-r)^{-m+a}$, provided $1 < a < 2$. For simplicity, we choose $a = (3/2)$.

Remark 4: Although we do not directly use his results, the relaxations in (II.13) and (II.14) are similar in spirit to Polak's theory of consistent approximations [36]. Also note that, even for a simple Euler discretization, appropriate relaxations of the endpoint condition and path constraint are essential to guarantee convergence [11].

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

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

where w_k are the LGL weights given by

$$w_k = \frac{2}{N(N+1)}\frac{1}{[L_N(t_k)]^2}$$

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

In utilizing existing nonlinear programming software, it is often necessary to provide a search region for the algorithm. For this reason, the following constraints are added:

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

where $\mathbb{X}$ and $\mathbb{U}$ are two compact sets presenting the search region and containing the continuous optimal solution $(x^*(t), u^*(t))$. Hence, the optimal control Problem B can be approximated to a nonlinear programming problem with $\bar{J}^N$ as the objective function and (II.11), (II.13), and (II.14) as constraints; this is summarized as follows.

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

$$\bar{J}^N(\bar{X}, \bar{U}) = \sum_{k=0}^{N} F\left(\bar{x}_k^N, \bar{u}_k^N\right) w_k + E\left(\bar{x}_0^N, \bar{x}_N^N\right) \tag{II.15}$$

subject to

$$D\begin{pmatrix} \bar{x}_{10}^N \\ \vdots \\ \bar{x}_{1N}^N \end{pmatrix} = \begin{pmatrix} \bar{x}_{20}^N \\ \vdots \\ \bar{x}_{2N}^N \end{pmatrix}$$
$$\vdots$$
$$D\begin{pmatrix} \bar{x}_{r-1,0}^N \\ \vdots \\ \bar{x}_{r-1,N}^N \end{pmatrix} = \begin{pmatrix} \bar{x}_{r0}^N \\ \vdots \\ \bar{x}_{rN}^N \end{pmatrix} \tag{II.16}$$
$$D\begin{pmatrix} \bar{x}_{r0}^N \\ \vdots \\ \bar{x}_{rN}^N \end{pmatrix} = \begin{pmatrix} f\left(\bar{x}_0^N\right) + g\left(\bar{x}_0^N\right)\bar{u}_0^N \\ \vdots \\ f\left(\bar{x}_N^N\right) + g\left(\bar{x}_N^N\right)\bar{u}_N^N \end{pmatrix}$$
$$\left\|e\left(\bar{x}_0^N, \bar{x}_N^N\right)\right\|_\infty \leq (N-r)^{-m+\frac{3}{2}} \tag{II.17}$$
$$h\left(\bar{x}_k^N, \bar{u}_k^N\right) \leq (N-r)^{-m+\frac{3}{2}} \cdot \mathbf{1}. \tag{II.18}$$

Remark 5: It is clear that the state equations in (II.16) have a lower triangular form. Decision variables $(\bar{x}_{20}^N, \ldots, \bar{x}_{2N}^N, \ldots, \bar{x}_{r0}^N, \ldots, \bar{x}_{rN}^N)$ are linear combinations of $(\bar{x}_{10}^N, \ldots, \bar{x}_{1N}^N)$. The preservation of the triangular structure can be exploited for computational efficiency as illustrated in [43].

III. Main Results

In the previous section, we formulated a pseudospectral method for solving continuous optimal control problems. In this approach, a continuous optimal control problem is approximated by a problem of nonlinear programming, which can be solved by an appropriate globally convergent algorithm [4], such as for example, a sequential-quadratic programming method. Although this approach has been successfully used in solving an impressive array of problems (see, for example, [43], [15], [20], [37], and [44]), some fundamental questions regarding the existence and convergence of the approximations have heretofore remained open. More specifically, we motivate and investigate the following questions.

- Is there a feasible solution to the nonlinear algebraic (II.16)–(II.17)–(II.18), if a feasible solution to the continuous optimal control problem exists?
- Under what condition does a sequence of feasible solutions of discrete Problem B^{N} converge to a feasible solution of continuous Problem B?
- If a sequence of discrete optimal solutions converges as the number of nodes increases, does it converge to the original optimal solution of Problem B?

These questions are not only important from a theoretical standpoint, but they are also important practical questions, particularly for real-time computation. In this section, we provide answers to these fundamental questions for feedback linearizable systems.

A. Existence of a Solution to Problem $\mathbf{B}^{\mathbf{N}}$

In the case of Eulerian discretizations, for any given initial condition and control series, the states are uniquely determined. Hence, there always exists a feasible solution to the discrete dynamic system. For Runge–Kutta methods, a similar property holds if the mesh is sufficiently dense [26]. For pseudospectral methods an existence result for controlled differential equations is not readily apparent. There are two main difficulties. PS methods are fundamentally different than traditional methods (like Euler or Runge–Kutta) 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 [5]. This technique cannot be used for controlled differential equations as it implies that the control can take arbitrary values at the boundary. Thus, PS methods for control differ from their standard counterparts in imposing the differential equation at the boundary as well. An unfortunate consequence of this approach is that the discretized dynamics may not have a feasible solution as illustrated by the following example.

Example 3: Consider linear system

$$\begin{aligned} \dot{x}_1 &= x_1 + u \\ \dot{x}_2 &= x_2 + u. \end{aligned} \tag{III.19}$$

Its PS discretization is

$$D\begin{pmatrix} \bar{x}_{10}^N \\ \vdots \\ \bar{x}_{1N}^N \end{pmatrix} = \begin{pmatrix} \bar{x}_{10}^N \\ \vdots \\ \bar{x}_{1N}^N \end{pmatrix} + \begin{pmatrix} \bar{u}_0^N \\ \vdots \\ \bar{u}_N^N \end{pmatrix}$$
$$D\begin{pmatrix} \bar{x}_{20}^N \\ \vdots \\ \bar{x}_{2N}^N \end{pmatrix} = \begin{pmatrix} \bar{x}_{20}^N \\ \vdots \\ \bar{x}_{2N}^N \end{pmatrix} + \begin{pmatrix} \bar{u}_0^N \\ \vdots \\ \bar{u}_N^N \end{pmatrix}.$$

Therefore

$$(D-I)\begin{pmatrix} \bar{x}_{10}^N \\ \vdots \\ \bar{x}_{1N}^N \end{pmatrix} = (D-I)\begin{pmatrix} \bar{x}_{20}^N \\ \vdots \\ \bar{x}_{2N}^N \end{pmatrix}.$$

Since D is nilpotent, $(D-I)$ is nonsingular. Hence, $(\bar{x}_{10}^N,\ldots,\bar{x}_{1N}^N) = (\bar{x}_{20}^N,\ldots,\bar{x}_{2N}^N)$. Therefore, if the initial condition is such that $\bar{x}_{10}^N \neq \bar{x}_{20}^N$, the discretized dynamics with arbitrary initial conditions has no solution, although a continuous solution satisfying (III.19) always exists for any given initial condition.

It can be shown that, for all uncontrollable linear systems, the discretized dynamical equations have no feasible solution for arbitrary initial conditions. The problem of existence of a solution is further exacerbated for nonlinear systems.

The infeasibility problem can be overcome by either restricting the system dynamic to some special structure like feedback linearizable systems considered in this paper, or simply relaxing the dynamics in the same manner as the relaxing of the end-point condition and path constraints [24], [25]. In Theorem I, we prove that the feasibility of problem B^{N} is guaranteed. In addition, the construction of a feasible solution developed in Theorem I is used in the proof of the convergence result in Theorem 2. However, first, we need the following lemma.

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

$$|\xi(t) - p^N(t)| \leq CC_0N^{-m} \qquad \forall\, t \in [-1,1] \tag{III.20}$$

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

Proof: This is a standard result of polynomial approximations; see [10]. ■

Theorem 1: Given any feasible solution, $t \mapsto (x,u)$, for Problem B. Suppose $x_r(\cdot) \in W^{m,\infty}$ 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, $(\bar{x}_k^N, \bar{u}_k^N), k = 0,\ldots,N$. Furthermore, the feasible solution satisfies

$$\left|x(t_k) - \bar{x}_k^N\right| \leq L(N-r)^{1-m} \tag{III.21}$$
$$\left|u(t_k) - \bar{u}_k^N\right| \leq L(N-r)^{1-m} \tag{III.22}$$

for all $k = 0,\ldots,N$, where L is a positive constant independent of N.

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

$$|\dot{x}_r(t) - p(t)| \le C_1(N-r)^{1-m} \qquad \forall t \in [-1,1]. \quad \text{(III.23)}$$

Define

$$\begin{aligned}
\hat{x}_r(t) &= \int_{-1}^{t} p(\tau)d\tau + x_r(-1) \\
\hat{x}_{r-1}(t) &= \int_{-1}^{t} \hat{x}_r(\tau)d\tau + x_{r-1}(-1) \\
&\vdots \\
\hat{x}_1(t) &= \int_{-1}^{t} \hat{x}_2(\tau)d\tau + x_1(-1) \\
\hat{u}(t) &= \frac{p(t) - f(\hat{x}_1(t),\ldots,\hat{x}_r(t))}{g(\hat{x}_1(t),\ldots,\hat{x}_r(t)).}
\end{aligned}$$

Clearly, $\hat{x}_1(t),\ldots,\hat{x}_r(t)$ are polynomials of degree less than or equal to N that satisfy the differential (II.3) with the initial condition, $\hat{x}(-1) = x(-1)$. By definition, the derivatives of a polynomial of degree less than or equal to N evaluated at the nodes $t_0,\ldots,t_N$ are exactly equal to the values of the polynomial at the nodes multiplied by the differentiation matrix D [10]. Thus, if we let

$$\bar{x}_k^N = \hat{x}(t_k) \quad \bar{u}_k^N = \hat{u}(t_k)$$

we have

$$\begin{aligned}
D\begin{pmatrix} \bar{x}_{i0}^N \\ \vdots \\ \bar{x}_{iN}^N \end{pmatrix} &= D\begin{pmatrix} \hat{x}_i(t_0) \\ \vdots \\ \hat{x}_i(t_N) \end{pmatrix} = \begin{pmatrix} \dot{\hat{x}}_i(t_0) \\ \vdots \\ \dot{\hat{x}}_i(t_N) \end{pmatrix} \\
&= \begin{pmatrix} \hat{x}_{i+1}(t_0) \\ \vdots \\ \hat{x}_{i+1}(t_N) \end{pmatrix} = \begin{pmatrix} \bar{x}_{i+1,0}^N \\ \vdots \\ \bar{x}_{i+1,N}^N \end{pmatrix}
\end{aligned}$$

where $i = 1,2,\ldots,r-1$ and $\bar{x}_{ik}^N$ is the ith component of $\bar{x}_k^N$. At $i = r$, we have

$$\begin{aligned}
D\begin{pmatrix} \bar{x}_{r0}^N \\ \vdots \\ \bar{x}_{rN}^N \end{pmatrix} &= D\begin{pmatrix} \hat{x}_r(t_0) \\ \vdots \\ \hat{x}_r(t_N) \end{pmatrix} = \begin{pmatrix} \dot{\hat{x}}_r(t_0) \\ \vdots \\ \dot{\hat{x}}_r(t_N) \end{pmatrix} \\
&= \begin{pmatrix} f(\hat{x}(t_0)) + g(\hat{x}(t_0))\hat{u}(t_0) \\ \vdots \\ f(\hat{x}(t_N)) + g(\hat{x}(t_N))\hat{u}(t_N) \end{pmatrix}.
\end{aligned}$$

Therefore, $(\bar{x}_k^N, \bar{u}_k^N), k = 0,1,\ldots,N$, satisfy the discrete differential constraints, (II.16).

Next, we prove that the mixed state-control constraints (II.18) are also satisfied. Based on (III.23)

$$\begin{aligned}
|x_r(t) - \hat{x}_r(t)| &\le 2C_1(N-r)^{1-m} \\
&\vdots \\
|x_1(t) - \hat{x}_1(t)| &\le 2^r C_1(N-r)^{1-m}.
\end{aligned}$$

Since $m \ge 2$ (by assumption), both $x(t)$ and $\hat{x}(t)$ are contained in some compact set whose boundary is independent of N. Furthermore, as f and g are Lipschitz continuous in the compact set (also by assumption), there is a constant C_2 independent of N such that

$$\begin{aligned}
|u(t) - \hat{u}(t)| = &\left| \frac{\dot{x}_r(t) - f(x_1(t),\ldots,x_r(t))}{g(x_1(t),\ldots,x_r(t))} - \right. \\
&\left. \frac{p(t) - f(\hat{x}_1(t),\ldots,\hat{x}_r(t))}{g(\hat{x}_1(t),\ldots,\hat{x}_r(t))} \right| \\
&\le C_2(|\dot{x}_r(t) - p(t)| + |x_1(t) - \hat{x}_1(t)| + \cdots \\
&+ |x_r(t) - \hat{x}_r(t)|).
\end{aligned}$$

Hence, we have, for $i = 1,\ldots,r$

$$|x_i(t) - \hat{x}_i(t)| \le C_3(N-r)^{1-m} \qquad \text{(III.24)}$$

$$|u(t) - \hat{u}(t)| \le C_3(N-r)^{1-m} \qquad \text{(III.25)}$$

for all $t \in [-1,1]$ and for some positive constant C_3 independent of N. Because $h(\cdot)$ is Lipschitz continuous, we can also write

$$\begin{aligned}
&\|h(x(t),u(t)) - h(\hat{x}(t),\hat{u}(t))\|_\infty \\
&\quad\le C_4(|x_1(t) - \hat{x}_1(t)| + \cdots + |x_r(t) - \hat{x}_r(t)| \\
&\qquad + |u(t) - \hat{u}(t)|) \\
&\quad\le C_3C_4(r+1)(N-r)^{1-m}
\end{aligned}$$

where C_4 is the Lipschitz constant of $h(\cdot)$ which is independent of N. Hence

$$\begin{aligned}
&h(\hat{x}(t),\hat{u}(t)) \\
&\quad\le h(x(t),u(t)) + C_3C_4(r+1)(N-r)^{1-m}\cdot\mathbf{1} \\
&\quad\le C_3C_4(r+1)(N-r)^{1-m}\cdot\mathbf{1}.
\end{aligned}$$

Since $m \ge 2$, there exists a positive integer N_1 such that, for all $N > N_1$,

$$C_3C_4(r+1)(N-r)^{1-m} \le (N-r)^{-m+\frac{3}{2}}.$$

Therefore, $\hat{x}_1(t_k),\ldots,\hat{x}_r(t_k),\hat{u}(t_k), k = 0,1,\ldots,N$, satisfy the mixed state and control constraint (II.18) for all $N > N_1$.

By a similar procedure, the fact that the endpoint condition is satisfied according to (II.17) can be proved. Hence, $(\bar{x}_k^N, \bar{u}_k^N)$ is a feasible discretized solution. At $t = t_k$, (III.24) and (III.25) imply (III.21) and (III.22). Thus, we have constructed a feasible solution to problem B^N that satisfies (III.21) and (III.22). ■

B. Convergence of $(\bar{x}_k^N, \bar{u}_k^N)$

In the previous section, we proved that if Problem B has a feasible solution, then Problem B^N also has a feasible solution. Here, we will show that a sequence of feasible solutions of Problem B^N converges to a feasible solution of Problem B as $N \to \infty$. In numerical analysis of ordinary differential equations, the convergence of Euler or Runge–Kutta methods are well known. Sufficient conditions under which the numerical solution converges to the continuous solution of an ordinary differential equation (ODE) can be found in textbooks. However, despite many successful applications of pseudospectral methods, few results can be found in the literature regarding the convergence of pseudospectral approximation of ordinary differential equations. In this section, we provide a mild and numerically checkable condition to guarantee such a convergence.

Let $(\bar{x}_k^N, \bar{u}_k^N), k = 0, 1, \ldots, N$, be a feasible solution of Problem $\mathrm{B^N}$. Denote $x_i^N(t)$ the Nth-order interpolating polynomial of $(\bar{x}_{i0}^N, \ldots, \bar{x}_{iN}^N)$, i.e.,

$$x_i^N(t) = \sum_{k=0}^{N} \bar{x}_{ik}^N \phi_k(t), \qquad i = 1, 2, \ldots, r \tag{III.26}$$

where $\phi_k(t)$ is defined by (II.7). Thus, $x_i^N(t_k) = \bar{x}_{ik}^N$ for all $i = 1, 2, \ldots r$ and $0 \le k \le N$. Also, denote

$$u^N(t) = \frac{\dot{x}_r^N(t) - f(x_1^N(t), \ldots, x_r^N(t))}{g(x_1^N(t), \ldots, x_r^N(t)).} \tag{III.27}$$

From (II.9) and (II.16), we have $u^N(t_k) = \bar{u}_k^N$ for all $0 \le k \le N$.

Now, consider a sequence of feasible solution of Problem $\mathrm{B^N}$ with N increasing from N_1 to infinity. Correspondingly, we get a sequence of interpolating functions $\{x_1^N(t), \ldots, x_r^N(t), u^N(t)\}_{N=N_1}^{\infty}$. Our convergence result is based on the following assumption.

Assumption 1: For each $1 \le i \le r$, the sequence $\{\bar{x}_{i0}^N\}_{N=N_1}^{\infty}$ converges as $N \to \infty$. Furthermore, there exists a continuous function $q(t)$ such that $\dot{x}_r^N(t)$ converges to $q(t)$ uniformly on $t \in [-1, 1]$.

Remark 6: In many optimal control problems, an initial value of the state is fixed by the endpoint condition. Then, from (II.17), it is easy to see the convergence of $\{\bar{x}_{i0}^N\}_{N=N_1}^{\infty}$ in Assumption 1 is automatically guaranteed. The second part of the assumption requires the convergence of the derivative of the interpolating polynomials of the last state variable. In the next section, we provide a practical method to verify this assumption.

Theorem 2: Let $\{(\bar{x}_k^N, \bar{u}_k^N), 0 \le k \le N\}_{N=N_1}^{\infty}$ be a sequence of feasible solution to Problem $\mathrm{B^N}$. Suppose Assumption 1 holds. Then, $\{\bar{x}_k^N, \bar{u}_k^N\}_{N=N_1}^{\infty}$ converges uniformly in k to a feasible solution of Problem B. More specifically, there exists a solution of the differential (II.3), $(x^\infty(t), u^\infty(t))$, satisfying the endpoint condition (II.4) and the path constraint (II.5), such that the following limit converges uniformly in k:

$$\lim_{N\to\infty} \left(\bar{x}_k^N - x^\infty(t_k)\right) = 0 \tag{III.28}$$
$$\lim_{N\to\infty} \left(\bar{u}_k^N - u^\infty(t_k)\right) = 0. \tag{III.29}$$

In addition, let $\{\bar{x}^N(t)\}_{N=N_1}^{\infty}$ be a sequence of polynomials defined by (III.26) and $\{\bar{u}^N(t)\}_{N=N_1}^{\infty}$ be a sequence of functions constructed by (III.27), then the following limit converges uniformly in t:

$$\lim_{N\to\infty} (\bar{x}^N(t) - x^\infty(t)) = 0 \tag{III.30}$$
$$\lim_{N\to\infty} (\bar{u}^N(t) - u^\infty(t)) = 0. \tag{III.31}$$

Proof: Based on Assumption 1, let $q(t)$ be the limit of $\dot{x}_r^N(t)$ and x_{i0} be the limit of $\{\bar{x}_{i0}^N\}_{N=N_1}^{\infty}$. Then, define the following functions:

$$x_r^\infty(t) = \int_{-1}^{t} q(\tau)d\tau + x_{r0}$$
$$x_{r-1}^\infty(t) = \int_{-1}^{t} x_r^\infty(\tau)d\tau + x_{r-1,0}$$
$$\vdots$$
$$x_1^\infty(t) = \int_{-1}^{t} x_2^\infty(\tau)d\tau + x_{10}$$
$$u^\infty(t) = \frac{q(t) - f(x_1^\infty(t), \ldots, x_r^\infty(t))}{g(x_1^\infty(t), \ldots, x_r^\infty(t)).}$$

Clearly, the pair, $(x^\infty(t), u^\infty(t))$, satisfies the differential (II.3). Let $x_i^N(t)$ be the interpolating polynomial of $\bar{x}_{i0}^N, \cdots, \bar{x}_{iN}^N$ defined by (III.26). Because $(\bar{x}_k^N, \bar{u}_k^N)$ satisfies discrete state (II.16), it is easy to see

$$\begin{pmatrix} \dot{x}_i^N(t_0) \\ \vdots \\ \dot{x}_i^N(t_N) \end{pmatrix} = D \begin{pmatrix} \bar{x}_{i0}^N \\ \vdots \\ \bar{x}_{iN}^N \end{pmatrix} = \begin{pmatrix} \bar{x}_{i+1,0}^N \\ \vdots \\ \bar{x}_{i+1,N}^N \end{pmatrix} = \begin{pmatrix} x_{i+1}^N(t_0) \\ \vdots \\ x_{i+1}^N(t_N) \end{pmatrix}$$

for $i = 1, 2, \ldots, r-1$. Hence the Nth-order polynomial,

$$\dot{x}_i^N(t) - x_{i+1}^N(t)$$

has $N+1$ distinct roots$t_0, \ldots, t_N$. Therefore, $\dot{x}_i^N(t) = x_{i+1}^N(t), i = 1, \ldots, r-1$. By Assumption 1, $\dot{x}_r^N(t)$ converges uniformly to $q(t)$; hence,

$$\lim_{N\to\infty} x_r^N(t) = \lim_{N\to\infty} \int_{-1}^{t} \dot{x}_r^N(\tau)d\tau + x_{r0} = \int_{-1}^{t} q(\tau)d\tau + x_{r0} = x_r^\infty(t).$$

Moreover, it is easy to show that this convergence is uniform in t. Therefore

$$\lim_{N\to\infty} x_{r-1}^N(t) = \lim_{N\to\infty} \int_{-1}^{t} x_r^N(\tau)d\tau + x_{r-1,0} = \int_{-1}^{t} x_r^\infty(\tau)d\tau + x_{r-1,0} = x_{r-1}^\infty(t).$$

Following the same procedure, we can prove

$$\lim_{N\to\infty} x_i^N(t) = x_i^\infty(t),\ i = 1, 2, \ldots, r \tag{III.32}$$
$$\lim_{N\to\infty} u^N(t) = u^\infty(t) \tag{III.33}$$

uniformly in t. Therefore, (III.28)–(III.31) are proved. The initial condition $e(x^\infty(-1), x^\infty(1)) = 0$ follows directly from the convergence property, since

$$e(x^\infty(-1), x^\infty(1)) = \lim_{N\to\infty} e(x^N(-1), x^N(1)) = \lim_{N\to\infty} e(\bar{x}_0^N, \bar{x}_N^N) = 0.$$

To prove that $(x^\infty(t), u^\infty(t))$ is a feasible solution to Problem B, it is now sufficient to prove that the mixed state-control constraint $h(x^\infty(t), u^\infty(t)) \le 0$ is satisfied. Suppose that at some time $t' \in (-1, 1]$ there is some $i \in [1, 2, \ldots, s]$ so that

$$h_i(x^\infty(t'), u^\infty(t')) > 0.$$

Since the nodes t_k are dense as N tends to infinity [21], there exists a sequence j^N that satisfies

$$0 \le j^N \le N$$
$$\lim_{N\to\infty} t_{j^N} = t'$$

Because (III.28) and (III.29) converge uniformly, we have

$$\lim_{N\to\infty} h_i(\bar{x}^N_{j^N}, \bar{u}^N_{j^N}) = h_i(x^\infty(t'), u^\infty(t')) > 0.$$

This contradicts the mixed state-control constraint, (II.18), in which the relaxation on the right side of the inequality approaches zero as $N \to \infty$. ■

C. Convergence of the Approximate Optimal Solutions

In the previous section, we proved a sufficient condition under which a sequence of discrete feasible solutions of Problem B^N converges to a feasible solution of the original continuous optimal control problem. Now, we study a sequence of special discrete feasible solutions. These are the optimal solutions of Problem B^N. Naturally, the question we must answer is: Under what condition does the sequence converge to the optimal solution of the continuous problem, and the cost (II.15) converges to the optimal cost function (II.2). At a first glance, the answer to this question seems simple if the numerical ODE solver used in the approximation is convergent. However, a close analysis reveals the difficulty when optimization is involved. Indeed, despite the simplicity of the Euler's method, the convergence of the Eulerian discretization of continuous optimal control problems has been an active research subject for a long period of time and a general error estimation was obtained just a few years ago [13]. Given the difficulty posed by Eulerian methods, it should be no surprise that proofs of convergence for non-Eulerian methods are even more difficult. Actually there are counter examples showing that a sequence of discrete optimal control based on certain Runge–Kutta approximation method does not converge to the optimal solution of the original problem [26], despite the fact that the system satisfies the standard conditions of Butcher. This is partially because dualization and discretization are not commutative operations [42] and, hence, standard convergence theorems associated with the discretization of differential equations are not applicable for the analysis of optimal control problems.

To prove the convergence of the discrete optimal solutions, existing results require strong conditions, for instance, the coercivity type of conditions [13], [26] or Lipschitz continuity of the inverse KKT mapping [24]. These conditions are not easily verifiable. Some conditions require information of the continuous optimal solution which is extremely difficult to get for nonlinear constrained systems. In the next section, we prove Theorem 3 to show that, under a practically verifiable condition (Assumption 1), a sequence of discrete optimal solutions converges to the optimal solution of the original continuous problem. Different from Theorem 2, where we prove the convergence of discrete *feasible* solutions, now the optimality of the solution is also guaranteed.

Theorem 3: Let $\{(\bar{x}^{*N}_k, \bar{u}^{*N}_k), 0 \le k \le N\}_{N=N_1}^\infty$ be a sequence of optimal solution of Problem B^N. Suppose the sequence satisfies Assumption 1. Then, there exists an optimal solution, $(x^*(\cdot), u^*(\cdot))$, to Problem B such that the following limits converge uniformly for $0 \le k \le N$

$$\lim_{N\to\infty} \left(\bar{x}^{*N}_k - x^*(t_k)\right) = 0$$
$$\lim_{N\to\infty} \left(\bar{u}^{*N}_k - u^*(t_k)\right) = 0$$
$$\lim_{N\to\infty} \bar{J}^N(\bar{X}^*, \bar{U}^*) = J(x^*(\cdot), u^*(\cdot)). \quad \text{(III.34)}$$

In order to prove Theorem 3, we need the following two lemmas.

Lemma 2: 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,

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

Proof: See [21]. ■

Lemma 3: Let $t_k, k = 0, 1, \ldots, N$, be LGL nodes. Suppose that $x(t)$ and $u(t)$ are continuous on $[-1, 1]$. Assume

$$\lim_{N\to\infty} |\bar{x}^N_k - x(t_k)| = 0 \quad \text{(III.35)}$$
$$\lim_{N\to\infty} |\bar{u}^N_k - u(t_k)| = 0 \quad \text{(III.36)}$$

uniformly in k, then we have

$$\lim_{N\to\infty} \bar{J}^N(\bar{X}, \bar{U}) = J(x(\cdot), u(\cdot)) \quad \text{(II.37)}$$

where $\bar{J}^N$ and J are the cost functions defined by (II.2) and (II.15), respectively.

Proof: From the uniform convergence property of $(\bar{x}^N_k, \bar{u}^N_k)$, it is easy to conclude $(\bar{x}^N_k, \bar{u}^N_k)$ is bounded for all $N \ge 1$ and $0 \le k \le N$. Therefore, by the fact that $F(x, u)$ is Lipschitz continuous, we have

$$\left|F(x(t_k), u(t_k)) - F\left(\bar{x}^N_k, \bar{u}^N_k\right)\right| \le K\left(\left|x(t_k) - \bar{x}^N_k\right| + \left|u(t_k) - \bar{u}^N_k\right|\right)$$

for some $K > 0$ and for all $N \ge 1, 0 \le k \le N$. Furthermore, $F(x(t), u(t))$ is continuous in t. Thus, by Lemma 2, we have

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

Therefore

$$\int_{-1}^{1} F(x(t),u(t))dt = \lim_{N\to\infty}\left(\sum_{k=0}^{N} F\left(\bar{x}_k^N,\bar{u}_k^N\right)w_k + \sum_{k=0}^{N}\left(F(x(t_k),u(t_k)) - F\left(\bar{x}_k^N,\bar{u}_k^N\right)\right)w_k\right).$$

From the uniform convergence of (III.35) and (III.36) and the property of w_k

$$\sum_{k=0}^{N} w_k = 2$$

we know that

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

Thus

$$\int_{-1}^{1} F(x(t),u(t))dt = \lim_{N\to\infty}\sum_{k=0}^{N} F\left(\bar{x}_k^N,\bar{u}_k^N\right)w_k. \quad \text{(III.38)}$$

It is obvious that

$$\lim_{N\to\infty} E\left(\bar{x}_0^N,\bar{x}_N^N\right) = E(x(-1),x(1)). \quad \text{(III.39)}$$

Thus, the limit in (III.37) follows (III.38) and (III.39). ■

Proof of Theorem 3: From Theorem 2, we know that the discrete optimal solutions converge uniformly to a feasible trajectory of the continuous problem. More specifically, there exists a continuous feasible solution, $(x^\infty(t),u^\infty(t))$, of Problem B such that

$$\begin{aligned}\lim_{N\to\infty}\left(\bar{x}_k^* - x^\infty(t_k)\right) = 0\\ \lim_{N\to\infty}\left(\bar{u}_k^* - u^\infty(t_k)\right) = 0\end{aligned}$$

uniformly for $0 \le k \le N$. Let, $\bar{J}^N(\bar{X}^*,\bar{U}^*)$ and $J(x^*(\cdot),u^*(\cdot))$ denote the optimal costs of Problem $\mathrm{B^N}$ and Problem B, respectively, i.e.,

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

where $(x^*(\cdot),u^*(\cdot))$ denotes any optimal solution of Problem B satisfying $x_r^*(\cdot)\in W^{m,\infty}$ with $m\ge 2$, (the optimal solution may not be unique). According to Theorem 1, there exists a sequence of feasible solutions, $(\tilde{x}_k^N,\tilde{u}_k^N)$, to Problem $\mathrm{B^N}$ that converge uniformly to $(x^*(t),u^*(t))$. Now, from Lemma 3 and the optimality of $(x^*(t),u^*(t))$ and $(\bar{x}_k^{*N},\bar{u}_k^{*N})$, we have

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

Hence, $J(x^*(\cdot),u^*(\cdot)) = J(x^\infty(\cdot),u^\infty(\cdot))$. This is equivalent to saying that $(x^\infty(t),u^\infty(t))$ is a feasible solution that achieves the optimal cost. Therefore, $(x^\infty(t),u^\infty(t))$ is an optimal solution to the continuous optimal control Problem B. ■

IV. Remarks and Extensions

A. Some Remarks on the Existence and Convergence

By Theorems 2 and 3, if the discrete optimal solution converges, the limit points must lie on the optimal solution of the continuous problem. Therefore, without knowing the optimal solution to the continuous optimal control Problem B, one can still verify the optimality and the feasibility of the discrete solution. The proof of the theorems also established a stronger result in which the interpolating polynomials $x^N(t)$ of the discrete solution uniformly converge to the continuous solution $x^*(t)$; and the corresponding $u^N(t)$ uniformly converges to $u^*(t)$ for $t\in[-1,1]$. Therefore, we can reconstruct the continuous optimal solution from the discrete one by interpolation. As a result, the optimal control can be evaluated at points different from the LGL nodes.

Note that, although we focus our attention on the Legendre PS method, Theorems 1 and 2 remain valid for other types of nodes and discretization based on interpolation such as the ones discussed in [16], [19]. More specifically, given any set of nodes, $\{t_k^N\}_{1\le k\le N,N\ge 1}$, of an interval $[a,b]$, suppose that the nodes are dense. Let $\phi_k^N(t)$ be the interpolating polynomial of order N, which equals 1 at t_k^N and 0 at all other nodes. Then, define the matrix D so that its ith row consists of the value of derivatives of $\phi_k^N(t)$ at t_i^N for $0\le i\le N$. Then, (II.16)–(II.18) is a discretization of Problem B. Similar proofs can be applied to show that Theorems 1 and 2 hold (under Assumption 1) for this set of nodes. There are essentially two reasons why we prefer to use LGL nodes in our algorithm: 1) it allows us to use accurate Gauss quadrature integration thereby improving the accuracy and convergence rate of the discrete approximation; and 2) LGL type of quadrature nodes can effectively prevent the Runge phenomenon [10], [5] which violates Assumption 1.

In the previous sections, we focused on single input nonlinear systems in feedback linearizable normal form. The extension of the existence and convergence results for multiple-input feedback linearizable normal forms, such as those discussed in [31], is trivial. For the purpose of brevity, we omit this part of the analysis. For general nonlinear systems, the necessary and sufficient conditions under which the system is transformable to the normal form can be found in [31]. Although pseudospectral methods can be applied to much general nonlinear systems, we recommend the use of feedback linearizable normal forms, whenever possible, for several reasons. Theoretically, a normal form facilities a proof of the existence and convergence results as discussed. Computationally, a normal form presents advantages in terms of run time and robustness, all other things being equal. Interested readers are referred to [43], where performance comparisons between different system representations are discussed.

Theorem 3 proves the convergence property, but does not give any information about the convergence rate. There are two factors that are critical to the rate of convergence. One is the smoothness of the functions in Problem B and its solution. The other is the rate at which $\dot{x}_r^N(t)$ converges in Assumption 1. In the case when both the optimal solution and the nonlinear system are smooth, the convergence of the discrete optimal solution is at the same rate as the sequence $\dot{x}_r^N(t)$ in Assumption 1. If $\dot{x}_r^N(t)$ converges to $q(t)$ superlinearly, then $(\bar{x}_k^*, \bar{u}_k^*)$ converges to $(x^*(t_k), u^*(t_k))$ superlinearly too. A study on the problem of convergence rate is important from both theoretical and practical viewpoints; however, it is outside the scope of this paper.

In this paper, we require the optimal solution of the continuous system belongs to $W^{m,\infty}$ with $m \geq 2$. Based on the Sobolev Imbedding Theorem, this assumption implies $x_r^*(t)$ be continuously differentiable, which in turn requires the optimal controller $u^*(t)$ be continuous. However, it is possible for an optimal control problem to admit a discontinuous solution, for example a bang–bang type of controller. The results presented in the paper can be generalized to the discontinuous case; see [32] for details.

B. A Numerical Method for Verifying Assumption 1

Since Assumption 1 was critical to the proofs of Theorems 2 and 3, it is important to devise a practical method to test its validity. This can be done by transforming the interpolating polynomials to spectral space [5], [25]. We discuss this technique for the Legendre PS method.

Let $y^N(t)$ be a polynomial of order N. Then, $y^N(t)$ can also be expressed as

$$y^N(t) = \sum_{k=0}^{N} a_k L_k(t)$$

where a_k are constant coefficients and $L_k(t)$ are Legendre polynomials. For a given sequence of polynomials $(y^0(t), y^1(t), \ldots, y^N(t))$, their Legendre expansions are defined by a matrix equation of the following form:

$$\begin{pmatrix} y^0 \\ y^1 \\ \vdots \\ y^N \end{pmatrix} = \begin{pmatrix} a_0^0 & 0 & 0 & \cdots & 0 \\ a_0^1 & a_1^1 & 0 & \cdots & 0 \\ & & \vdots & & \\ a_0^N & a_1^N & \cdots & a_{N-1}^N & a_N^N \end{pmatrix} \begin{pmatrix} L_0(t) \\ L_1(t) \\ \vdots \\ L_N(t) \end{pmatrix}.$$

The convergence property of the polynomial sequence $\{y^N(t)\}$ can be characterized by the convergence of their Legendre coefficients $(a_k^0, a_k^1, \ldots, a_k^N), \qquad k = 0, \ldots, N$.

Lemma 4: Suppose the spectral coefficients $a_k^N, k = 1, 2, \ldots,$ converge as $N \to \infty$; and satisfy the following inequality:

$$|a_k^N - a_k^*| \leq cN^{-\beta}$$

where $a_k^*, c > 0$ and $\beta > 1$ are constants independent of N. Moreover, assume $\sum_{k=0}^{N} a_k^*$ converges absolutely. Then, $y^N(t)$ converges uniformly to $y(t) = \sum_{k=0}^{\infty} a_k^* L_k(t)$ on $t \in [-1, 1]$.

The proof of this result is straightforward. It is omitted here.

Once an optimal solution to Problem B^{N} is computed, the spectral coefficients of $\dot{x}^N(t)$ can be easily calculated by a matrix multiplication [5]

$$\begin{pmatrix} a_0 \\ \vdots \\ a_{N-1} \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & & \\ & \ddots & \\ & & N - \frac{1}{2} \end{pmatrix} \times \begin{pmatrix} L_0(t_0) & \cdots & L_0(t_N) \\ & \vdots & \\ L_{N-1}(t_0) & \cdots & L_{N-1}(t_N) \end{pmatrix} \times \begin{pmatrix} w_0 & & \\ & \ddots & \\ & & w_N \end{pmatrix} \times D \times \begin{pmatrix} \bar{x}_{r0}^* \\ \vdots \\ \bar{x}_{rN}^* \end{pmatrix}.$$

Thus, Assumption 1 can be numerically verified by examining the convergence of these coefficients [25]. Although this procedure is not equivalent to a mathematical verification of Assumption 1, we note that similar techniques are frequently adopted in practical scientific computation.

V. Examples

In this section, we present several examples to illustrate the main points of the PS method. All problems were programmed in MATLAB by way of DIDO [38] running on a Pentium 4, 2.4-GHz PC with 256 MB of RAM.

Example 4: Consider the one-link flexible robot arm discussed in the text [31] and in [43]. The system is modelled by

$$\begin{aligned} I_1\ddot{q}_1 + m_1 gl \sin q_1 + k(q_1 - q_2) &= 0 \\ I_2\ddot{q}_2 - k(q_1 - q_2) &= u \end{aligned}$$

where q_1, q_2 are the angular positions and $I_1 = I_2 = k = 1, g = 9.8, m_1 = 0.01$ and $l = 0.5$. The optimal control problem is to minimize

$$\int_0^1 u^2(t)dt$$

subject to endpoint constraint

$$\begin{aligned} [q_1(0), \dot{q}_1(0), q_2(0), \dot{q}_2(0)] &= [0.03, 0.04, 0.01, 0.05] \\ [q_1(1), \dot{q}_1(1), q_2(1), \dot{q}_2(1)] &= [0.06, 0.08, 0.02, 0.02]. \end{aligned}$$

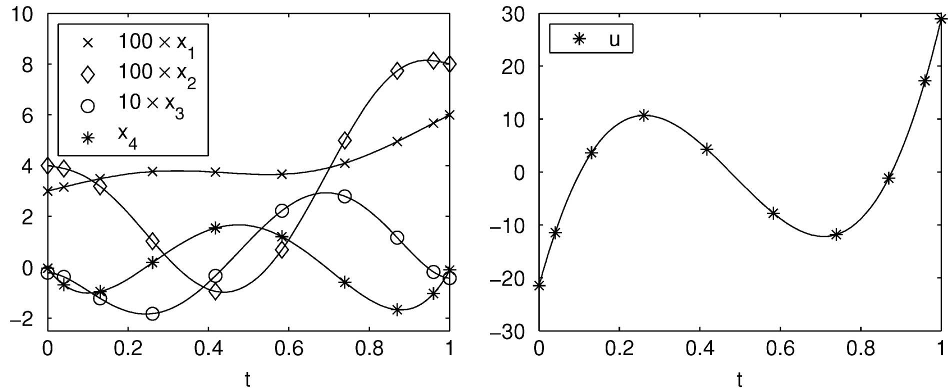


Fig. 5. Discrete optimal solution for Example 4 with ten nodes. The solid lines are generated using 100 nodes.

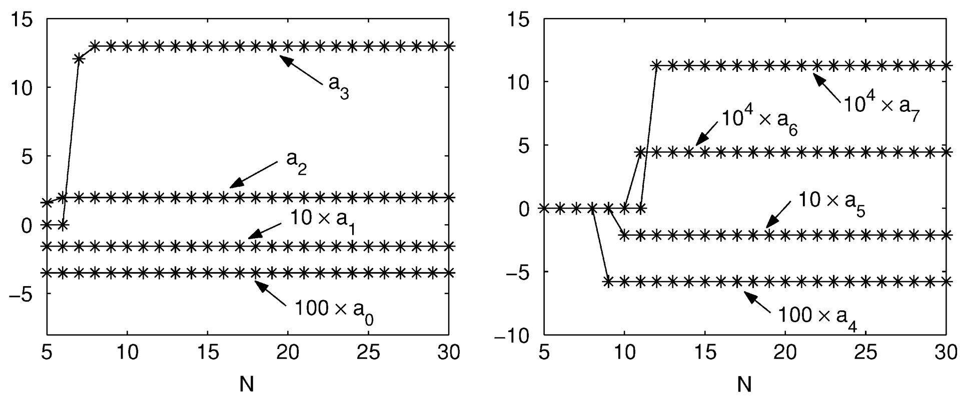


Fig. 6. Convergence of the spectral coefficients for Example 4.

By the transformation

$$\begin{aligned} x_1 &= q_1 \\ x_2 &= \dot{q}_1 \\ x_3 &= -\frac{1}{I_1}[m_1 g l \sin q_1 + k(q_1 - q_2)] \\ x_4 &= -\frac{1}{I_1}[m_1 g l \dot{q}_1 \cos q_1 + k(\dot{q}_1 - \dot{q}_2)] \end{aligned}$$

the system can be easily put into a normal form

$$\begin{aligned} \dot{x}_1 &= x_2 \\ \dot{x}_2 &= x_3 \\ \dot{x}_3 &= x_4 \\ \dot{x}_4 &= (\beta_2 \cos x_1 + \beta_3)x_3 + (\beta_4 x_2^2 + \beta_5)\sin x_1 + \beta_1 u \end{aligned}$$

where $\beta_1 = (k/I_1 I_2), \beta_2 = (-m_1 g l/I_1), \beta_3 = -(k(I_1 + I_2)/I_1 I_2), \beta_4 = -\beta_2$ and $\beta_5 = -(m_1 k g l/I_1 I_2)$. The solution for $N = 10$ is plotted in Fig. 5. The analytic solution to this problem is unavailable. However, the convergence of the computed solution can still be numerically verified by the method described in Section IV-B. To this end, we compute the spectral coefficients a_k of the polynomial,

$$\dot{x}_4^N(t) = \sum_{k=0}^{N} a_k L_k(t)$$

The results for $a_k, k = 0, \ldots, 7$ are shown in Fig. 6 for $N = 5, \ldots, 30$. It is apparent that we have a very fast convergence rate. Thus, all the conditions in Lemma 4 can be numerically verified; therefore, Assumption 1 holds. By Theorem 3, the convergence of the discrete solution is guaranteed, i.e.,

$$\lim_{N\to\infty} (\bar{x}_k^{*N}, \bar{u}_k^{*N}) = (x^*(t_k), u^*(t_k))$$

although an analytic expression of $(x^*(t_k), u^*(t_k))$ is unknown. In Fig. 5, we also plot the numerical solution for 100 nodes. Due to the fast convergence rate demonstrated in Fig. 6, the solution with 100 nodes can be reasonably treated as the "true" continuous optimal solution. Then, from Fig. 5, it is clear that a ten-node solution is sufficiently accurate for all practical purposes.

Example 5: Consider the Breakwell problem from [8]. The problem is to minimize

$$\frac{1}{2}\int_0^1 u^2(t)dt$$

subject to the differential equations

$$\dot{x}_1 = x_2 \quad \dot{x}_2 = u$$

the endpoint conditions

$$(x_1(0), x_2(0)) = (0, 1) \quad (x_1(1), x_2(1)) = (0, -1)$$

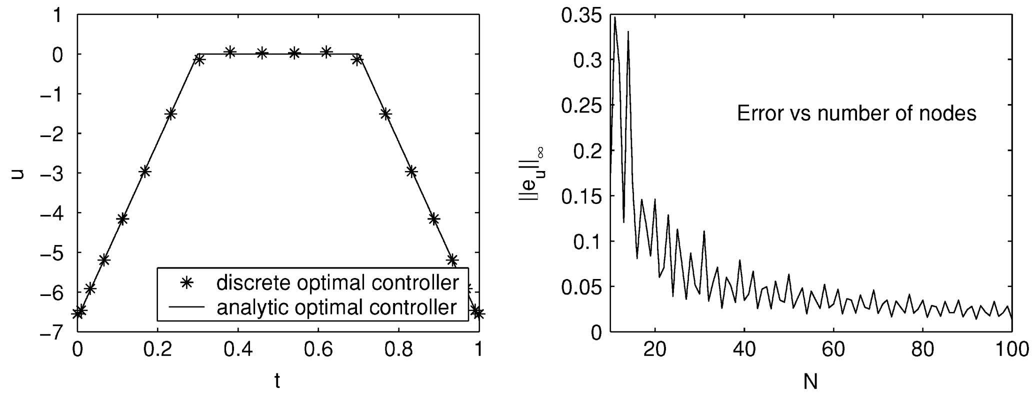


Fig. 7. Simulation results for the Breakwell problem.

and the state constraint

$$x_1(t) \leq 0.1.$$

The optimal control is given by [8]

$$u^*(t) = \begin{cases} \frac{200}{9}t - \frac{20}{3}, & t \in [0, 0.3] \\ 0, & t \in [0.3, 0.7] \\ -\frac{200}{9}t + \frac{140}{9}, & t \in [0.7, 1]. \end{cases}$$

The result for $\mathrm{N} = 20$ is shown in Fig. 7. We also plot the maximum error between the discrete and continuous control, i.e., $||e_u||_\infty = \max\{|\bar{u}_k^{*N} - u^*(t_k)|, k = 0, 1, \ldots, N\}$, in Fig. 7 for N ranging from 10 to 100. It can be seen that the error converges as N tends to infinity.

It is interesting to note that the costates are discontinuous [8], [42] for this problem. Hence, a proof of convergence by dualizing the problem is quite difficult even for an Euler discretization. On the other hand, by validating Assumption 1, it is not difficult to show the convergence of the discrete optimal solution despite the discontinuity in the costates.

Fig. 7 demonstrates the convergence of the discrete optimal controller. However, the convergence rate is not as impressive as previous examples. This is due to the lack of smoothness of the optimal solution. As discussed in Section IV-A, the convergence rate of PS methods depends highly on the smoothness of the solution. For this example, the optimal controller is only continuous; therefore, we expect a slow convergence rate. To address this and other issues, PS knotting methods have been developed in [41]. By using the concept of fixed soft knots [41], the error can be reduced to 10^{-6} with just 12 nodes. Although the convergence rate is not the focus of this paper, we note this point simply to emphasize an important issue that deserves further investigation.

VI. Conclusion

Although pseudospectral methods have been successfully applied to solve a wide variety of complex engineering problems, a rigorous proof of convergence of these methods has heretofore been unavailable. The difficulties arise because convergence proofs for optimal control problems belie intuition. As recent studies show, convergent Runge–Kutta methods may diverge while nonconvergent implicit Runge–Kutta (IRK) methods may converge. Thus, a new type of analysis is necessary to address these problems, which is why even Eulerian methods continue to be studied to this day. In this paper, we provide a theoretical foundation for the convergence of solutions obtained by PS methods for a class of constrained nonlinear systems. As our proof did not require dualizing the problem, we circumvented the difficulties associated with the convergence of discontinuous costates.

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. Comput.*, vol. 23, no. 6, pp. 1981–2007, 2002.

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

[3] J. T. Betts, "Survey of numerical methods for trajectory optimization," *J. Guid., Control, Dyna.*, vol. 21, no. 2, pp. 193–207, 1998.

[4] P. T. Boggs, A. J. Kearsley, and J. W. Tolle, "A global convergence analysis of an algorithm for large-scale nonlinear optimization problems," *SIAM J. Optim.*, vol. 9, no. 4, pp. 833–862, 1999.

[5] J. P. Boyd, *Chebyshev and Fourier Spectral Methods*, 2nd ed. New York: Dover, 2001.

[6] B. Bonnard, L. Faubourg, and E. Trélat, "Optimal control of the atmospheric arc of a space shuttle and numerical simulations by multiple-shotting techniques," *Math. Models Meth. Appl. Sci.*, vol. 2, no. 15, 2005.

[7] A. E. Bryson, *Dynamic Optimization*. Reading, MA: Addison-Wesley, 1999.

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

[9] J. B. Caillau and J. Noailles, "Coplanar control of a satellite around the Earth," *ESAIM Cont. Opt. Calc.*, vol. 6, pp. 239–258, 2001.

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

[11] J. Cullum, "Finite-dimensional approximations of state constrainted continuous optimal problems," *SIAM J. Control*, vol. 10, pp. 649–670, 1972.

[12] A. L. Dontchev, "Discrete approximations in optimal control," in *Nonsmooth Analysis and Geometric Methods in Deterministic Optimal Control, IMA Vol. Math. Appl.*. New York: Springer-Verlag, 1996, vol. 78, pp. 59–81, MR 97h:49043.

[13] A. L. Dontchev and W. W. Hager, "The Euler approximation in state constrained optimal control," *Math. Comput.*, vol. 70, pp. 173–203, 2000.

[14] A. L. Dontchev, W. W. Hager, and V. M. Veliov, "Second-order Runge-Kutta approximations in control constrained optimal control," *SIAM J. Numer. Anal.*, vol. 38, no. 1, pp. 202–226, 2000.

[15] G. Elnagar, M. A. Kazemi, and M. Razzaghi, "The pseudospectral Legendre method for discretizing optimal control problems," *IEEE Trans. Autom. Control*, vol. 40, no. 10, pp. 1793–1796, Oct. 1995.

[16] G. Elnagar and M. A. Kazemi, "Pseudospectral Chebyshev optimal control of constrained nonlinear dynamical systems," *Comput. Optim. Appl.*, vol. 11, pp. 195–217, 1998.

[17] F. Fahroo and I. M. Ross, "Costate estimation by a Legendre pseudospectral method," in *Proc. AIAA Guidance, Navigation and Control Conf.*, Boston, MA, Aug. 10–12, 1998, paper no. AIAA 98-4222.

[18] F. Fahroo and I. M. Ross, "Computational optimal control by spectral collocation with differential inclusion," in *Proc. 1999 Goddard Flight Mechanics Symp.*, pp. 185–200, NASA/CP-1999-209235.

[19] F. Fahroo and I. M. Ross, "Direct trajectory optimization by a Chebyshev pseudospectral method," *J. Guid., Control, Dyna.*, vol. 25, no. 1, pp. 160–166, 2002.

[20] F. Fahroo and I. M. Ross, "Costate estimation by a Legendre pseudospectral method," *J. Guid., Control, Dyna.*, vol. 24, no. 2, pp. 270–277, 2001.

[21] G. Freud, *Orthogonal Polynomials*. New York: Pergamon, 1971.

[22] P. E. Gill, W. Murray, and M. A. Saunders, "SNOPT: An SQP algorithm for large-scale constrained optimization," *SIAM J. Optim.*, vol. 12, no. 4, pp. 979–1006, 2002.

[23] C. J. Goh and K. L. Teo, "Control parametrization: A unified approach to optimal control problems with general constraints," *Automatica*, vol. 24, pp. 3–18, 1988.

[24] Q. Gong, I. M. Ross, W. Kang, and F. Fahroo, "Convergence of pseudospectral methods for constrained nonlinear optimal control problems," in *Proc. IASTED Int. Conf. Intelligent Systems and Control*, Honolulu, HI, 2004, pp. 209–214.

[25] Q. Gong and I. M. Ross, "Autonomous pseudospectral knotting methods," in *Proc. 16th AAS/AIAA Space Flight Mechanics Conf.*, Tampa, FL, Jan. 2006, vol. AAS 06-151.

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

[27] C. R. Hargraves and S. W. Paris, "Direct tracjectory optimization using nonlinear programming and collocation," *AIAA J. Guid., Control, Dyna.*, vol. 10, no. 4, pp. 338–342, 1987.

[28] R. K. Mehra and R. E. Davis, "A generalized gradient method for optimal control problems with inequality constraints and singular arcs," *IEEE Trans. Autom. Control*, vol. AC-17, no. 1, pp. 67–78, Jan. 1972.

[29] B. S. Mordukhovich, *Approximation Methods in Problems of Optimization and Control*. Moscow, Russia: Nauka, 1988.

[30] S. I. Infeld and W. Murray, "Optimization of stationkeeping for a libration point mission," in *Proc. AAS Spaceflight Mechanics Meeting*, Maui, HI, Feb. 2004, vol. AAS 04-150.

[31] A. Isidori, *Nonlinear Control Systems*. New York: Springer-Verlag, 1995.

[32] W. Kang, Q. Gong, and I. M. Ross, "Convergence of pseudospectral methods for nonlinear optimal control problems with discontinuous controller," in *Proc. 44th IEEE Conf. Decision and Control and European Control Conf. (CDC-ECC'05)*, Seville, Spain, 2005, pp. 2799–2804.

[33] P. Lu, H. Sun, and B. Tsai, "Closed-loop endoatmospheric ascent guidance," *J. Guid., Control, Dyna.*, vol. 26, no. 2, pp. 283–294, 2003.

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

[35] E. Polak, "A historical survey of computations methods in optimal control," *SIAM Rev.*, vol. 15, pp. 553–548, 1973.

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

[37] J. Rea, "Launch vehicle trajectory optimization using a Legendre pseudospectral method," in *Proc. AIAA Guidance, Navigation, Control Conf.*, Austin, TX, Aug. 2003, Paper No. AIAA 2003-5640.

[38] I. M. Ross, User's Manual for DIDO: A MATLAB application package for solving optimal control problems Tomlab Optimization, Inc., Feb. 2004, Tech. Rep. 04-01.0.

[39] I. M. Ross, "A historical introduction to the covector mapping principle," *Adv. Astronaut. Sci.: Astrodyna.*, vol. 122, Aug. 2005, Paper AAS 05-332.

[40] I. M. Ross and F. Fahroo, "Pseudospectral methods for optimal motion planning of differentially flat systems," *IEEE Trans. Autom. Control*, vol. 49, no. 8, pp. 1410–1413, Aug. 2004.

[41] I. M. Ross and F. Fahroo, "Pseudospectral knotting methods for solving optimal control problems," *AIAA J. Guid., Control, Dyna.*, vol. 27, no. 3, pp. 397–405, 2004.

[42] I. M. Ross and F. Fahroo, *Legendre Pseudospectral Approximations of Optimal Control Problems*, ser. Lecture Notes in Control and Information Sciences. New York: Springer-Verlag, 2003, vol. 295, pp. 327–342.

[43] I. M. Ross and F. Fahroo, "Issues in the real-time computation of optimal control," in *Mathematical and Computer Modelling: An International Journal*. New York: Elsevier, vol. 43.

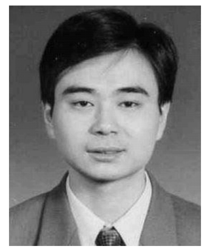
[44] S. Stanton, R. Proulx, and C. D. Souza, "Optimal orbit transfer using a Legendre pseudospectral method," in *Proc. AAS/AIAA Astrodynamics Specialist Conf.*, Big Sky, MT, Aug. 3–7, 2003, no. AAS-03-574.

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

[46] R. J. Vanderbei, "Case studies in trajectory optimization: Trains, planes and other pastimes," *Optim. Eng.*, vol. 2, pp. 215–243, 2001.

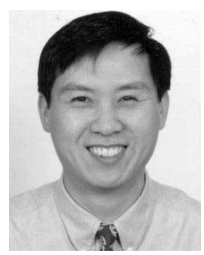
[47] R. Vinter, *Optimal Control*. Boston, MA: Birkhäuser, 2000.

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



Qi Gong (S'02–M'04) received the B.E. degree from Dalian University of Science and Technology, Dalian, China, the M.E. degree Southeast University, China, and the Ph.D. degree in electrical engineering and computer science from Case Western Reserve University, Cleveland, OH, in 1996, 1999, and 2004, respectively.

From March 2004 to March 2005, he was a National Research Council (NRC) Postdoctoral Researcher at the U.S. Naval Postgraduate School (NPS), Monterey, CA. He is currently a Research Associate in the Department of Mechanical and Astronautical Engineering at NPS. His research interests include computational optimal control and estimation, nonlinear control and observer design, robust and adaptive control, model predictive control, and their industry applications.



Wei Kang (M'91) received the B.S. and M.S. degrees from Nankai University, China, both in mathematics, in 1982 and 1985, respectively, and the Ph.D. degree in mathematics from the University of California, Davis, in 1991.

He is currently a Professor of applied mathematics at the U.S. Naval Postgraduate School (NPS), Monterey, CA. From 1991 to 1994, he was a Visiting Assistant Professor of systems science and mathematics at Washington University, St. Louis, MO. His research interest includes PS methods of optimal control, nonlinear systems with bifurcations, normal forms and invariants, cooperative control of autonomous vehicles, industry applications of control theory, nonlinear filtering, and nonlinear H-infinity control. His early research includes topics on Lie groups, Lie algebras, and differential geometry.

Dr. Kang received the Best Paper Award of the 6th International Conference on Control, Automation, Robotics, and Vision held in Singapore in 2000. He was the 2003 recipient of the Carl E. and Jessie W. Menneken Faculty Award of NPS Foundation.

I. Michael Ross received the Ph.D. degree in Aerospace Engineering from the Pennsylvania State University, University Park, in 1991.

Currently, he is a Professor and Program Director of Control and Optimization in the Department of Mechanical and Astronautical Engineering at the Naval Postgraduate School (NPS), Monterey, CA. He is the Principal Investigator on a pseudospectral feedback control experiment onboard NPSAT1, an experimental satellite scheduled to be launched in September 2006. He was the Project Lead on PANSAT, a small communications satellite currently in low-Earth-orbit. From 1999 to 2001, he was a Visiting Associate Professor at Draper Labs, Cambridge, MA, during which time he introduced pseudospectral methods for solving nonlinear control problems in the areas of launch and entry guidance, attitude control and inertial navigation. He is the author of DIDO, a popular software package for solving optimal control problems. His research interests are in real-time computation of optimal control with applications to nonlinear feedback control problems arising in spacecraft formation design and reconfiguration control, low-thrust orbital maneuvers, aerobraking, spacecraft singularity-free control, missile guidance, and space maneuvers for future systems.

Dr. Ross is an Associate Fellow of the AIAA and the Founding Book Review Editor for the *Journal of Guidance, Control and Dynamics*.