Direct Trajectory Optimization and Costate Estimation of Finite-Horizon and Infinite-Horizon Optimal Control Problems Using a Radau Pseudospectral Method

Divya Garg*
Michael A. Patterson†
Camila Francolin‡
Christopher L. Darby§
Department of Mechanical and Aerospace Engineering
University of Florida
Gainesville, FL 32611

Geoffrey T. Huntington¶
Blue Origin, LLC.
Kent, WA 98032

William W. Hager‖
Department of Mathematics
University of Florida
Gainesville, FL 32611

Anil V. Rao**
Department of Mechanical and Aerospace Engineering
University of Florida
Gainesville, FL 32611

*Ph.D. Student, Dept. of Mechanical and Aerospace Engineering. E-mail: divyagarg2002@ufl.edu
†Ph.D. Student, Dept. of Mechanical and Aerospace Engineering. E-mail: mpatterson@ufl.edu
‡Ph.D. Student, Dept. of Mechanical and Aerospace Engineering. E-mail: francoli@ufl.edu
§Ph.D. Student, Dept. of Mechanical and Aerospace Engineering. E-mail: cdarby@ufl.edu
¶GN& C Engineer. E-mail: ghuntington@blue.aero
‖Professor. E-mail: hager@math.ufl.edu
Assistant Professor. E-mail: anilvrao@ufl.edu. **Corresponding Author.

Abstract

A method is presented for direct trajectory optimization and costate estimation of finite-horizon and infinite-horizon optimal control problems using global collocation at Legendre-Gauss-Radau (LGR) points. A key feature of the method is that it provides an accurate way to map the KKT multipliers of the nonlinear programming problem to the costates of the optimal control problem. More precisely, it is shown that the dual multipliers for the discrete scheme correspond to a pseudospectral approximation of the adjoint equation using polynomials one degree smaller than that used for the state equation. The relationship between the coefficients of the pseudospectral scheme for the state equation and for the adjoint equation is established. Also, it is shown that the inverse of the pseudospectral LGR differentiation matrix is precisely the matrix associated with an implicit LGR integration scheme. Hence, the method presented in this paper can be thought of as either a global implicit integration method or a pseudospectral method. Numerical results show that the use of LGR collocation as described in this paper leads to the ability to determine accurate primal and dual solutions for both finite and infinite-horizon optimal control problems.

1 Introduction

Over the last decade, pseudospectral methods have increased in popularity in the numerical solution of optimal control problems.[1–16] Pseudospectral methods are a class of *direct collocation* where the optimal control problem is transcribed to a nonlinear programming problem (NLP) by parameterizing the state and control using global polynomials and collocating the differential-algebraic equations using nodes obtained from a Gaussian quadrature. It is noted that some researchers prefer the term *orthogonal collocation*,[17–19] but the terms pseudospectral and orthogonal collocation have the same meaning.

The three most commonly used set of collocation points are *Legendre-Gauss* (LG), *Legendre-Gauss-Radau* (LGR), and *Legendre-Gauss-Lobatto* (LGL) points. These three sets of points are obtained from the roots of a Legendre polynomial and/or linear combinations of a Legendre polynomial and its derivatives. All three sets of points are defined on the domain $[-1, 1]$, but differ significantly in that the LG points include *neither* of the endpoints, the LGR points include *one* of the endpoints, and the LGL points include *both* of the endpoints. In addition, the LGR points are asymmetric relative to the origin and are not unique in that they can be defined using either the initial point or the terminal point. In recent years, the two most well documented pseudospectral methods are the *Legendre-Gauss-Lobatto pseudospectral method*[1,3–5,10,11,20,21] (LPM) and the *Legendre-Gauss pseudospectral method*.[13–15,22,23] A *local collocation* method based on LGR points is developed in Ref. 16 while an LGR method

for solving *infinite-horizon* problems is in Ref. 12. Nonetheless, Legendre-Gauss-Radau collocation still remains the least studied of the pseudospectral methods. The purpose of this paper is to describe a method for direct trajectory optimization and costate estimation for general *finite-horizon and infinite-horizon* optimal control problems using global collocation at LGR points.

The pseudospectral LGR scheme presented in this paper is related to the scheme of Kameswaran and Biegler in Ref. 16, however, the focus of our paper is quite different. In Ref 16, the authors partition the time interval into a mesh and use the LGR scheme on each mesh interval. Convergence is achieved by increasing the number of mesh intervals. Here we focus on a global LGR scheme where convergence is achieved by increasing the degree of the polynomials. In Ref. 12 Fahroo and Ross apply an LGR scheme to infinite horizon control problems. This is done by using a change of variables to map the infinite time interval into a finite time interval, and then applying an LGR scheme to the finite time interval problem. Since the change of variables is singular at the final time, an LGR scheme avoided collocation at the singularity. As explained in Section 8, the LGR scheme introduced in Ref. 12 is fundamentally different from the LGR scheme presented here because in the method presented here the state is discretized at the LGR points *plus* the terminal point, thus allowing for the solution of *both* finite-horizon and infinite-horizon optimal control problems.

The approach developed in this paper is well-suited to problems that are sufficiently smooth; that is, problems that have neither discontinuities in the control nor large derivatives in the state. More generally, when approximating the solution to a control problem, a high quality approximation in a low dimensional space might be achieved by using piecewise polynomials with a high degree in time intervals where the solution is smooth, and with a low degree in time intervals where the solution lacks smoothness. In time intervals where the solution undergoes rapid change, the mesh could be refined to improve the accuracy of the approximation. As a first step towards developing a convergence theory for this framework where the degree of the piecewise polynomials is allowed to vary, we need to understand the convergence properties of discrete approximations generated by polynomials on a single interval, as the degree of the polynomials increase. As can be seen in Refs. 24 or 25, the key step in analyzing the conference of discrete approximations is to reformulate the first-order optimality conditions in such a way that they become an approximation to the optimality conditions for the continuous control problem. In this paper, we develop and analyze these

reformulated optimality conditions by exploiting a transformed adjoint variable.[26]

The paper is organized as follows: In Section 2 we discuss the choices for collocation points and our notation. Section 3 introduces our Radau pseudospectral scheme for an unconstrained control problem. In Section 4 we show that the first-order optimality conditions associated with the pseudospectral scheme can be written as a pseudospectral scheme for the adjoint equation. The polynomials associated with the transformed adjoint equation have degree one smaller than that of the polynomials associated with the state equation discretization. In Section 5 we show that our pseudospectral scheme is equivalent to an integrated system of equations. A modification of the method for infinite-horizon problems is then discussed in Section 6. Section 8 compares our scheme to the methods presented in Ref. 12 and 16. Finally, Sections 7 and 9 give numerical examples and conclusions.

2 LG, LGR, and LGL Collocation Points

The LG, LGR, and LGL collocation points lie on the open interval $\tau \in (-1, 1)$, the half open interval $\tau \in [-1, 1)$ or $\tau \in (-1, 1]$, and the closed interval $\tau \in [-1, 1]$, respectively. A depiction of these three sets of collocation points is shown in Fig. 1 where it is seen that the LG points contain neither -1 or 1, the LGR points contain only *one* of the points -1 or 1 (in this case, the point -1), and the LGL points contain *both* -1 and 1. Denoting K as the number of collocation points and $P_K(\tau)$ as the k^{th}-degree Legendre polynomial, the LG points are the roots of $P_K(\tau)$, the LGR points are the roots of $P_{K-1}(\tau) + P_K(\tau)$, and the LGL points are the roots of $\dot{P}_{K-1}(\tau)$ together with the points -1 and 1:

LG:	Roots obtained from $P_K(\tau)$
LGR:	Roots obtained from $P_{K-1}(\tau) + P_K(\tau)$
LGL:	Roots obtained from $\dot{P}_{k-1}(\tau)$ together with the points -1 and 1

It is seen from Fig. 1 that the LG and LGL points are symmetric about the origin whereas the LGR points are asymmetric. In addition, the LGR points are not unique in that two sets of points exist (one including the point -1 and the other including the point 1). The LGR points that include the terminal endpoint are often called the *flipped* LGR points. In this paper, however, we use the standard set of LGR points as defined above and consistent with the usage given in Ref. 20.

Notation. Throughout the paper, we employ the following notation. $\mathbf{A}^{\mathsf{T}}$ denotes the transpose of a matrix $\mathbf{A}$. Given two matrices $\mathbf{A}$ and $\mathbf{B}$ of the same dimensions, $\langle \mathbf{A}, \mathbf{B} \rangle$ is their dot product:

$$\langle \mathbf{A}, \mathbf{B} \rangle = \text{trace } \mathbf{A}^{\mathsf{T}}\mathbf{B}.$$

When $\mathbf{A}$ and $\mathbf{B}$ are vectors, this is the usual vector inner product. If $\mathbf{f} : \mathbb{R}^n \to \mathbb{R}^m$, then $\nabla \mathbf{f}$ is the m by n matrix whose i-th row is ∇f_i. In particular, the gradient of a scalar-valued function is a row vector. If $\phi : \mathbb{R}^{m \times n} \to \mathbb{R}$ and $\mathbf{X}$ is an m by n matrix, then $\nabla \phi$ denotes the m by n matrix whose (i, j) element is

$$(\nabla \phi(\mathbf{X}))_{ij} = \frac{\partial \phi(\mathbf{X})}{\partial X_{ij}}.$$

3 Formulation of Pseudospectral Method Using LGR Points

To simplify the exposition, we initially focus on an unconstrained control problem on the time interval $\tau \in [-1, +1]$ with terminal cost. Note that the time interval can be transformed from $[-1, 1]$ to the time interval $[t_0, t_f]$ via the affine transformation

$$t = \frac{t_f - t_0}{2}\tau + \frac{t_f + t_0}{2}.$$

In this section, the goal is to determine the state $\mathbf{x}(\tau) \in \mathbb{R}^n$ and the control $\mathbf{u}(\tau) \in \mathbb{R}^m$ which minimize the cost functional

$$\Phi(\mathbf{x}(1)) \tag{1}$$

subject to the constraints

$$\frac{d\mathbf{x}}{d\tau} = \mathbf{f}(\mathbf{x}(\tau), \mathbf{u}(\tau)), \quad \mathbf{x}(-1) = \mathbf{x}_0, \tag{2}$$

where $\mathbf{f} : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$ and $\mathbf{x}_0$ is the initial condition, which we assume is given.

Let us consider N LGR collocation points τ_1, τ_2, ..., τ_N on the interval $[-1, 1]$, with $\tau_1 = -1$ and $\tau_N < +1$. We introduce an additional noncollocated point $\tau_{N+1} = 1$ which is used to describe the approximation to the state variable. Each component of the state $\mathbf{x}$ is approximated by a polynomial of degree at most N. Let L_i, $i = 1, \ldots, N + 1$, be a basis of Lagrange polynomials given by

$$L_i(\tau) = \prod_{\substack{j=1 \\ j \neq i}}^{N+1} \frac{\tau - \tau_j}{\tau_i - \tau_j}, \qquad i = 1, \ldots, N + 1.$$

The j-th component of the state is approximated by a series of the form

$$x_j(\tau) \approx \sum_{i=1}^{N+1} x_{ij} L_i(\tau). \tag{3}$$

Differentiating the series and evaluating at the collocation point τ_k gives

$$\dot{x}_j(\tau_k) \approx \sum_{i=1}^{N+1} x_{ij} \dot{L}_i(\tau_k) = \sum_{i=1}^{N+1} D_{ki} x_{ij}, \quad D_{ki} = \dot{L}_i(\tau_k). \tag{4}$$

The N by $N+1$ matrix $\mathbf{D}$ is called the differentiation matrix. It has one row for each collocation point; the elements in the i^{th} column are the derivatives of the Lagrange polynomials evaluated at each of the collocation points. Let $\mathbf{X}$ denote the matrix formed from the coefficients x_{ij} in (3). With this notation, $\mathbf{DX}$ is an N by n matrix and (4) can be written

$$\dot{x}_j(\tau_i) \approx (\mathbf{DX})_{ij}.$$

This approximation is exact if the components of $\mathbf{x}$ are polynomials of degree at most N.

Let $\mathbf{U}$ be an N by m matrix with u_{ij} denoting the discrete approximation to the j-th component of the control evaluated at the i-th collocation point:

$$u_{ij} \approx u_j(\tau_i).$$

For the matrices $\mathbf{X}$ and $\mathbf{U}$, and later for the matrices of Lagrange multipliers, subscripts are used to denote rows of the matrix. In other words, $\mathbf{X}_i$ is the i-th row of $\mathbf{X}$. This row contains the components of the discrete approximation to $\mathbf{x}^{\mathsf{T}}(\tau_i)$. Let $\mathbf{F}(\mathbf{X}, \mathbf{U})$ denote an N by n matrix whose (i, j) element is given by

$$F_{ij}(\mathbf{X}, \mathbf{U}) = f_j(\mathbf{X}_i, \mathbf{U}_i), \quad 1 \le i \le N, \quad 1 \le j \le n. \tag{5}$$

In the pseudospectral approach, it is required that the system dynamics is satisfied at each of the N collocation points. With our notation, the discrete optimization problem takes the form

$$\text{minimize} \quad \Phi(\mathbf{X}_{N+1}) \quad \text{subject to} \quad \mathbf{DX} = \mathbf{F}(\mathbf{X}, \mathbf{U}), \quad \mathbf{X}_1 = \mathbf{x}_0, \tag{6}$$

where $\mathbf{x}_0$ is treated as a row vector. The optimization problem in (6) is a nonlinear programming problem.

We now develop the first-order optimality conditions for (6), also called as the KKT conditions of the NLP. The system dynamics in (6) is composed of Nn equations. Let $\mathbf{\Lambda}$

denote the N by n matrix of Lagrange multipliers associated with the system dynamics, and let $\boldsymbol{\mu}$ be a 1 by n row vector of Lagrange multipliers associated with the initial condition. The Lagrangian associated with (6) is

$$\mathcal{L}(\boldsymbol{\Lambda}, \mathbf{X}, \mathbf{U}) = \Phi(\mathbf{X}_{N+1}) + \langle \boldsymbol{\Lambda}, \mathbf{F}(\mathbf{X}, \mathbf{U}) - \mathbf{D}\mathbf{X} \rangle + \langle \boldsymbol{\mu}, \mathbf{x}_0 - \mathbf{X}_1 \rangle.$$

The KKT conditions of the NLP are obtained by differentiating $\mathcal{L}$ with respect to each component of $\mathbf{X}$ and $\mathbf{U}$. Since i ranges between 1 and N in (5), $\mathbf{F}(\mathbf{X}, \mathbf{U})$ is independent of $\mathbf{X}_{N+1}$. Differentiating the Lagrangian with respect to $\mathbf{X}_{N+1}$ gives us the condition

$$\nabla \Phi(\mathbf{X}_{N+1}) = \mathbf{D}_{N+1}^{\mathsf{T}} \boldsymbol{\Lambda}, \tag{7}$$

where $\mathbf{D}_j$ denotes the j-th *column* of $\mathbf{D}$. Differentiating the Lagrangian with respect to $\mathbf{X}_j$ gives

$$\sum_{i=1}^{N} D_{ij} \boldsymbol{\Lambda}_i = \boldsymbol{\Lambda}_j \nabla_X \mathbf{f}(\mathbf{X}_j, \mathbf{U}_j), \quad 2 \leq j \leq N. \tag{8}$$

Finally, differentiating with respect to $\mathbf{X}_1$ yield

$$\sum_{i=1}^{N} D_{i1} \boldsymbol{\Lambda}_i = \boldsymbol{\Lambda}_1 \nabla_X \mathbf{f}(\mathbf{X}_1, \mathbf{U}_1) - \boldsymbol{\mu}. \tag{9}$$

Differentiating with respect to the control $\mathbf{U}_j$, $1 \leq j \leq N$, gives

$$\boldsymbol{\Lambda}_j \nabla_U \mathbf{f}(\mathbf{X}_j, \mathbf{U}_j) = \mathbf{0}. \tag{10}$$

Let $\mathbf{D}_{j:k}$ denote the submatrix of $\mathbf{D}$ formed by columns j through k, and let $\mathbf{X}_{j:k}$ be the submatrix of $\mathbf{X}$ corresponding to *rows* j through k. The system dynamics in (6) can be rewritten

$$\mathbf{D}_{2:N+1} \mathbf{X}_{2:N+1} = \mathbf{F}(\mathbf{X}, \mathbf{U}) - \mathbf{D}_1 \mathbf{x}_0 \tag{11}$$

Similarly, the costate equations (8) and (9) can be rewritten

$$\mathbf{D}_{1:N}^{\mathsf{T}} \boldsymbol{\Lambda} = \nabla_X \langle \boldsymbol{\Lambda}, \mathbf{F}(\mathbf{X}, \mathbf{U}) \rangle - \mathbf{e}_1 \boldsymbol{\mu}, \tag{12}$$

where $\mathbf{e}_1$ is the first column of the identity matrix. We now observe that the N by N matrices appearing on the left sides of these equation are invertible.

Proposition 1. *The matrices* $\mathbf{D}_{1:N}$ *and* $\mathbf{D}_{2:N+1}$ *obtained by deleting either the first or the last column of* $\mathbf{D}$ *are invertible.*

Proof. Suppose that for some nonzero $\mathbf{p} \in \mathbb{R}^{N+1}$ with $p_{N+1} = 0$, we have $\mathbf{Dp} = \mathbf{0}$. Let p be the unique polynomial of degree N which satisfies $p(\tau_i) = p_i$, $1 \leq i \leq N+1$. Since the components of $\mathbf{Dp}$ are the derivatives of p evaluated at the collocation points, we have

$$0 = (\mathbf{Dp})_i = \dot{p}(\tau_i), \quad 1 \leq i \leq N.$$

Since $\dot{p}$ is a polynomial of $N-1$, it must be identically zero since it vanishes at N points. Hence, p is constant. Since $p(1) = 0$ and p is constant, it follows that p is identically 0. This shows that $p_i = p(\tau_i) = 0$ for each i. Since the equation $\mathbf{Dp} = \mathbf{0}$ with $p_{N+1} = 0$ has no nonzero solution, $\mathbf{D}_{1:N}$ is nonsingular. The nonsingularity of $\mathbf{D}_{2:N+1}$ is established in a similar way, but with $p_1 = 0$ instead of $p_{N+1} = 0$. □

4 Transformed Adjoint System

Analogous to Ref. 26, we now reformulate the KKT conditions of the NLP so that they become a discretization of the first-order optimality conditions for the continuous control problem (1)–(2). Let w_i, $1 \leq i \leq N$, be the quadrature weights associated with the LGR points. These quadrature weights have the property that

$$\int_{-1}^{1} p(\tau) d\tau = \sum_{i=1}^{N} w_i p(\tau_i)$$

for all polynomials p of degree at most $2N-2$. Let $\mathbf{W}$ denote the N by N diagonal matrix with i-th diagonal element w_i and let $\boldsymbol{\lambda}$ be an N by n matrix defined by

$$\boldsymbol{\lambda} = \mathbf{W}^{-1}\boldsymbol{\Lambda}. \tag{13}$$

We also define the row vector

$$\boldsymbol{\lambda}_{N+1} = \mathbf{D}_{N+1}^{\mathsf{T}}\boldsymbol{\Lambda}. \tag{14}$$

In the formulas that follow, it is convenient to consider $\boldsymbol{\lambda}$ as an N by n matrix and to view $\boldsymbol{\lambda}_{N+1}$ as a distinct row vector, not the $N+1$-st row of $\boldsymbol{\lambda}$. As we will see, the rows of $\boldsymbol{\lambda}$ as well as $\boldsymbol{\lambda}_{N+1}$ represent approximations to the continuous costate evaluated at τ_i, $1 \leq i \leq N+1$. In order to connect the discrete costate equations to the continuous costate equations, we employ an N by N matrix $\mathbf{D}^{\dagger}$, which is a modified version of $\mathbf{D}$, defined as follows:

$$D_{11}^{\dagger} = -D_{11} - \frac{1}{w_1} \quad \text{and} \quad D_{ij}^{\dagger} = -\frac{w_j}{w_i} D_{ji} \quad \text{otherwise.} \tag{15}$$

According to the definition of $\boldsymbol{\lambda}_{N+1}$, the adjoint boundary condition (7) is simply

$$\nabla\Phi(\mathbf{X}_{N+1}) = \boldsymbol{\lambda}_{N+1}. \tag{16}$$

Utilizing (13) and (15), (8) reduces to

$$\sum_{j=1}^{N} D_{ij}^{\dagger}\boldsymbol{\lambda}_j = -\boldsymbol{\lambda}_i\nabla_X\mathbf{f}(\mathbf{X}_i,\mathbf{U}_i), \quad 2 \le i \le N. \tag{17}$$

Similarly, (9) reduces to

$$\sum_{j=1}^{N} D_{1j}^{\dagger}\boldsymbol{\lambda}_j = -\boldsymbol{\lambda}_1\nabla_X\mathbf{f}(\mathbf{X}_1,\mathbf{U}_1) + \frac{1}{w_1}(\boldsymbol{\mu} - \boldsymbol{\lambda}_1). \tag{18}$$

Finally, dividing (10) by w_j yields

$$\boldsymbol{\lambda}_i\nabla_U\mathbf{f}(\mathbf{X}_i,\mathbf{U}_i) = \mathbf{0}, \quad 1 \le i \le N. \tag{19}$$

The equations (16)–(19) are incomplete since we introduced a new variable $\boldsymbol{\lambda}_{N+1}$ without adding a new equation. We now develop an equation for this new variable by manipulating (14). Let $\mathbf{1}$ denote a vector whose components are all equal to 1. The components of the vector $\mathbf{D1}$ are the derivatives at the collocation points of the polynomial whose value is 1 at τ_i, $1 \le i \le N+1$. This polynomial is simply the constant 1, whose derivative is 0 everywhere. Hence, we have $\mathbf{D1} = \mathbf{0}$, which implies that

$$\mathbf{D}_{N+1} = -\sum_{j=1}^{N}\mathbf{D}_j. \tag{20}$$

Returning to the definition of $\boldsymbol{\lambda}_{N+1}$ in (14), we obtain

$$\begin{aligned}
\boldsymbol{\lambda}_{N+1} &= \sum_{i=1}^{N}\boldsymbol{\Lambda}_i D_{i,N+1} = -\sum_{i=1}^{N}\sum_{j=1}^{N}\boldsymbol{\Lambda}_i D_{ij} && (21)\\
&= \frac{\boldsymbol{\Lambda}_1}{w_1} + \sum_{i=1}^{N}\sum_{j=1}^{N}\boldsymbol{\Lambda}_i D_{ji}^{\dagger}\frac{w_j}{w_i} = \frac{\boldsymbol{\Lambda}_1}{w_1} + \sum_{i=1}^{N}\sum_{j=1}^{N}\boldsymbol{\Lambda}_j D_{ij}^{\dagger}\frac{w_i}{w_j} && (22)\\
&= \boldsymbol{\lambda}_1 + \sum_{i=1}^{N}\sum_{j=1}^{N} w_i\boldsymbol{\lambda}_j D_{ij}^{\dagger} && (23)\\
&= \boldsymbol{\mu} - \sum_{i=1}^{N} w_i\boldsymbol{\lambda}_i\nabla_X\mathbf{f}(\mathbf{X}_i,\mathbf{U}_i), && (24)
\end{aligned}$$

where (21) follows from the identity (20), (22) is the definition (15) of $\mathbf{D}^\dagger$, (23) is the definition (13) of $\boldsymbol{\lambda}_i$, and (24) is the first-order optimality condition (17). Together (16)–(19) and (24) form the complete transformed KKT conditions. More compactly, the KKT conditions are

$$\begin{aligned} \boldsymbol{\mu} &= \nabla\Phi(\mathbf{X}_{N+1}) + \sum_{i=1}^{N} w_i \boldsymbol{\lambda}_i \nabla_X \mathbf{f}(\mathbf{X}_i, \mathbf{U}_i), && (25) \\ \mathbf{D}^\dagger \boldsymbol{\lambda} &= -\nabla_X \langle \boldsymbol{\lambda}, \mathbf{F}(\mathbf{X}, \mathbf{U}) \rangle + \frac{1}{w_1} \mathbf{e}_1 (\boldsymbol{\mu} - \boldsymbol{\lambda}_1), && (26) \\ \mathbf{0} &= \nabla_U \langle \boldsymbol{\lambda}, \mathbf{F}(\mathbf{X}, \mathbf{U}) \rangle. \end{aligned}$$

where (25) is obtained by combining (16) and (24).

We now compare the transformed KKT conditions for the discrete control problem (the pseudospectral scheme) to the first-order optimality condition for the continuous control problem (1)–(2):

$$\begin{aligned} \boldsymbol{\lambda}(-1) &= \boldsymbol{\mu} \\ \boldsymbol{\lambda}(1) &= \nabla\Phi(\mathbf{x}(1)) \\ \dot{\boldsymbol{\lambda}}(t) &= -\nabla_x \langle \boldsymbol{\lambda}(t), \mathbf{f}(\mathbf{x}(t), \mathbf{u}(t)) \rangle \\ \mathbf{0} &= \nabla_u \langle \boldsymbol{\lambda}(t), \mathbf{f}(\mathbf{x}(t), \mathbf{u}(t)) \rangle \end{aligned}$$

In the discrete problem, there is no multiplier corresponding to the final time $\tau = 1$ since the system dynamics are collocated at τ_i, $1 \le i \le N$, which are all strictly less than 1. In the discrete optimality system, the boundary conditions for the continuous optimality system are replaced by the integrated version (25). $\nabla\Phi(\mathbf{X}_{N+1})$ in (25) corresponds to $\boldsymbol{\lambda}(1)$ in the continuous problem; the summation in (25) approximates the integral of $\dot{\boldsymbol{\lambda}}$ over the interval $[-1, 1]$. Hence, the right side of (25) approximates $\boldsymbol{\lambda}(-1)$, which corresponds to $\boldsymbol{\lambda}_1$. Consequently, the condition (25) is a subtle way of enforcing the equality $\mu = \boldsymbol{\lambda}_1$, in an approximate sense. If $\mu = \boldsymbol{\lambda}_1$, then last term in the discrete dynamics (26) vanishes. We will now show that the system (26), with the last term dropped, is a pseudospectral scheme for the costate equation.

Theorem 1. *The matrix $\mathbf{D}^\dagger$ defined in* (15) *is the differentiation matrix for the space of polynomials of degree $N-1$ evaluated at τ_i, $1 \le i \le N$. In other words, if p is a polynomial of degree at most $N-1$ and if $\mathbf{p} \in \mathbb{R}^N$ is the vector with i-th component $p_i = p(\tau_i)$, then*

$$(\mathbf{D}^\dagger \mathbf{p})_i = \dot{p}(\tau_i), \quad 1 \le i \le N.$$

Proof. Let $\mathbf{D}^\dagger$ denote the differentiation matrix defined in the statement of the theorem. We will show that $\mathbf{D}^\dagger$ satisfies (15), which establishes the theorem. If p and q are smooth, real-valued functions with $p(1) = 0$, then integration by parts gives

$$\int_{-1}^{1} \dot{p}(\tau)q(\tau)d\tau = -p(-1)q(-1) - \int_{-1}^{1} p(\tau)\dot{q}(\tau)d\tau. \tag{27}$$

Suppose p is a polynomial of degree at most N and q is a polynomial of degree at most $N-1$ with $N \geq 1$; in this case, $\dot{p}q$ and $p\dot{q}$ are polynomials of degree at most $2N-2$. Since Gauss-Radau quadrature is exact for polynomials of degree at most $2N-2$, the integrals in (27) can be replaced by their quadrature equivalents to obtain

$$\sum_{j=1}^{N} w_j \dot{p}_j q_j = -p_1 q_1 - \sum_{j=1}^{N} w_j p_j \dot{q}_j,$$

where $p_j = p(\tau_j)$ and $\dot{p}_j = \dot{p}(\tau_j)$. More compactly, this can be expressed

$$(\mathbf{W}\dot{\mathbf{p}})^\mathsf{T}\mathbf{q} = -p_1 q_1 - (\mathbf{W}\mathbf{p})^\mathsf{T}\dot{\mathbf{q}}.$$

Substituting $\dot{\mathbf{p}} = \mathbf{D}_{1:N}\mathbf{p}$ and $\dot{\mathbf{q}} = \mathbf{D}^\dagger\mathbf{q}$ yields

$$\mathbf{p}^\mathsf{T}\mathbf{D}_{1:N}^\mathsf{T}\mathbf{W}\mathbf{q} = -p_1 q_1 - \mathbf{p}^\mathsf{T}\mathbf{W}\mathbf{D}^\dagger\mathbf{q}.$$

This can be rearranged into the following form:

$$\mathbf{p}^\mathsf{T}(\mathbf{D}_{1:N}^\mathsf{T}\mathbf{W} + \mathbf{W}\mathbf{D}^\dagger + \mathbf{e}_1\mathbf{e}_1^\mathsf{T})\mathbf{q} = 0.$$

where $\mathbf{e}_1$ is the first column of the identity matrix. Since this identity must be satisfied for all choices of $\mathbf{p}$ and $\mathbf{q}$, we deduce that

$$\mathbf{D}_{1:N}^\mathsf{T}\mathbf{W} + \mathbf{W}\mathbf{D}^\dagger + \mathbf{e}_1\mathbf{e}_1^\mathsf{T} = \mathbf{0},$$

which implies (15). This completes the proof. □

Thus we have shown that the transformed KKT conditions are related to a pseudospectral discretization of the continuous costate equation. However, the differentiation matrix in the costate discretization is based on the derivatives of polynomials of degree $N-1$, while the differentiation matrix in the state discretization is based on the derivatives of polynomials of degree N.

5 Integral Formulation

Next, we show that pseudospectral discretization of the state equation has an equivalent integrated formulation. Similar to (20), the identity $\mathbf{D1} = \mathbf{0}$ implies that

$$\mathbf{D}_1 = -\sum_{j=2}^{N+1} \mathbf{D}_j = -\mathbf{D}_{2:N+1}\mathbf{1}. \tag{28}$$

By Proposition 1 the matrix $\mathbf{D}_{2:N+1}$ is invertible. We multiply (28) by $\mathbf{D}_{2:N+1}^{-1}$ to obtain

$$\mathbf{D}_{2:N+1}^{-1}\mathbf{D}_1 = -\mathbf{1}. \tag{29}$$

Let p be any polynomial of degree at most N. By the construction of the N by $N+1$ differentiation matrix $\mathbf{D}$, we have $\mathbf{Dp} = \dot{\mathbf{p}}$ where

$$p_i = p(\tau_i), \quad 1 \le i \le N+1 \quad \text{and} \quad \dot{p}_i = \dot{p}(\tau_i), \quad 1 \le i \le N. \tag{30}$$

Multiply the identity $\dot{\mathbf{p}} = \mathbf{Dp} = \mathbf{D}_1 p_1 + \mathbf{D}_{2:N+1}\mathbf{p}_{2:N+1}$ by $\mathbf{D}_{2:N+1}^{-1}$ and utilize (29) to obtain

$$p_i = p_1 + (\mathbf{D}_{2:N+1}^{-1}\dot{\mathbf{p}})_i, \quad 2 \le i \le N+1. \tag{31}$$

Next, we obtain a different expression for $p_i - p_1$ based on the integration of the interpolant of the derivative. Let $L_j^\dagger$ be the Lagrange interpolation polynomial associated with the collocation points:

$$L_j^\dagger(\tau) = \prod_{\substack{i=1 \\ i \ne j}}^{N} \frac{\tau - \tau_i}{\tau_j - \tau_i}, \qquad j = 1, \ldots, N.$$

Given a polynomial p of degree at most N, its derivative $\dot{p}$ is a polynomial of degree at most $N-1$. Hence, $\dot{p}$ can be interpolated exactly by the Lagrange polynomials $L_j^\dagger$:

$$\dot{p}(\tau) = \sum_{j=1}^{N} \dot{p}_j L_j^\dagger(\tau)$$

We integrate from -1 to τ_i to obtain the relation

$$p(\tau_i) = p(-1) + \sum_{j=1}^{N} \dot{p}_j A_{ij}, \quad A_{ij} = \int_{-1}^{\tau_i} L_j^\dagger(\tau)\, d\tau, \quad 2 \le i \le N+1. \tag{32}$$

Utilizing the notation (30), we have

$$p_i = p_1 + (\mathbf{A}\dot{\mathbf{p}})_i, \quad 2 \le i \le N+1. \tag{33}$$

The relations (31) and (33) are satisfied for any polynomial of degree at most N. Choose $p_1 = 0$ and $\dot{\mathbf{p}}$ from the columns of the identity matrix to deduce that $\mathbf{A} = \mathbf{D}_{2:N}^{-1}$. Multiply (11) by $\mathbf{A} = \mathbf{D}_{2:N}^{-1}$ and utilize (29) to obtain

$$\mathbf{X}_i = \mathbf{x}_0 + \mathbf{A}_i \mathbf{F}(\mathbf{X}, \mathbf{U}), \quad 2 \leq i \leq N+1, \tag{34}$$

where $\mathbf{A}_i$ is the i^{th} row of $\mathbf{A}$. Hence, the differential form of the state equation $\mathbf{DX} = \mathbf{F}(\mathbf{X}, \mathbf{U})$ is equivalent to the integrated form (34), where the elements of $\mathbf{A}$ are integrals of the Lagrange basis functions $L_j^{\dagger}$ defined in (32) while the elements of $\mathbf{D}$ in the differential form are the derivatives of the Lagrange basis function L_i defined in (4).

To summarize, the approximation to the dynamics given in (34) is in the form of a global *implicit integration method* while the differential approximation $\mathbf{DX} = \mathbf{F}(\mathbf{X}, \mathbf{U})$ is in the form of a *pseudospectral method*. The fact that either the integral or the differential form can be used shows that the Radau collocation method derived in this paper can be thought of as *either* a global implicit integration method *or* a pseudospectral method. In particular, using the pseudospectral form of LGR collocation results in a system of equations that has *no loss* of information from the integral form (because the matrix $\mathbf{D}_{2:N}$ is *nonsingular*). In the next section, the differential form of LGR collocation, which we call the *Radau pseudospectral method*, is applied to a general optimal control problem.

6 Radau Pseudospectral Discretization of Infinite-Horizon Problems

Consider the following optimal control problem. Minimize the infinite-horizon cost functional

$$J = \int_0^{\infty} g(\mathbf{x}(t), \mathbf{u}(t), t) dt \tag{35}$$

subject to the dynamic constraint

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}(t), \mathbf{u}(t), t) \tag{36}$$

with the initial condition

$$\mathbf{x}(0) = \mathbf{x}_0. \tag{37}$$

Consider further the following transformation of time found in Ref. 12:

$$t = \frac{1+\tau}{1-\tau} \tag{38}$$

This transformation maps the interval $t \in [0, \infty)$ to the closed interval $\tau \in [-1, 1]$. Using (38), the infinite-horizon optimal control problem (35)–(37) can be written in terms of τ as follows. Minimize the cost functional

$$J = \int_{-1}^{1} \frac{2}{(1-\tau)^2} g(\mathbf{x}(\tau), \mathbf{u}(\tau), \tau) d\tau \tag{39}$$

subject to the dynamic constraint

$$\frac{d\mathbf{x}}{d\tau} = \frac{2}{(1-\tau)^2} \mathbf{f}(\mathbf{x}(\tau), \mathbf{u}(\tau), \tau) \tag{40}$$

with the initial condition

$$\mathbf{x}(-1) = \mathbf{x}_0 \tag{41}$$

The transformed infinite-horizon optimal control problem (39)–(41) can be solved using the following modification of the Radau pseudospectral discretization of Section 3. Minimize the cost function

$$J = \sum_{k=1}^{N} \frac{2w_k}{(1-\tau_k)^2} g(\mathbf{X}_k, \mathbf{U}_k, \tau_k) \tag{42}$$

subject to the constraints

$$\begin{aligned} \mathbf{DX} &= \mathbf{TF}(\mathbf{X}, \mathbf{U}) \\ \mathbf{X}_1 &= \mathbf{x}_0 \end{aligned} \tag{43}$$

where $\mathbf{T}$ is a diagonal matrix whose k^{th} diagonal element is

$$\mathbf{T}_{kk} = \frac{2}{(1-\tau_k)^2}, \quad 1 \leq k \leq N. \tag{44}$$

It is noted in the NLP of (42)–(43) that the state is approximated at the LGR points *plus* the terminal point (at $\tau = 1$). Hence we obtain an approximation of the state at the horizon $t = \infty$. Moreover, the NLP avoids the singularity at $\tau = +1$ in the factor $2/(1-\tau)^2$ because $\tau_k = +1$ is not a quadrature point. As is discussed in Section 8 below, the solution obtained using the Radau pseudospectral method of this paper differs fundamentally from the infinite-horizon method given in Ref. 12 because in the method of Ref. 12 the state is obtained only at the LGR points whereas in the method presented here the state is obtained at the LGR points *and* the terminal point $\tau = +1$.

7 Examples

In this section we consider two examples using the aforementioned Radau pseudospectral method. The first example is a nonlinear one-dimensional finite-horizon optimal control

problem taken from Ref. 13 while the second example is an infinite-horizon linear quadratic problem taken from Ref. 12. It is noted that these two examples utilize the finite-horizon and infinite-horizon forms of the Radau pseudospectral method, respectively.

Example 1: Nonlinear One-Dimensional Finite-Horizon Problem

Consider the following optimal control problem. Minimize the cost functional

$$J = \frac{1}{2}\int_0^{t_f} (y + u^2)dt \tag{45}$$

subject to the dynamic constraint

$$\dot{y} = 2y + 2u\sqrt{y}, \tag{46}$$

and the boundary conditions

$$\begin{aligned} y(0) &= 2, \\ y(t_f) &= 1, \\ t_f &= 5. \end{aligned} \tag{47}$$

It is noted that the exact solution to the optimal control problem of (45)–(47) is given as

$$\begin{aligned} y^*(t) &= x^2(t) \\ \lambda_y^*(t) &= \frac{\lambda_x}{2\sqrt{y}} \end{aligned} \tag{48}$$

where $x(t)$ and $\lambda_x(t)$ are given as

$$\begin{bmatrix} x(t) \\ \lambda_x(t) \end{bmatrix} = \exp(\mathbf{A}t) \begin{bmatrix} x_0 \\ \lambda_{x0} \end{bmatrix} \tag{49}$$

where

$$\begin{aligned} \mathbf{A} &= \begin{bmatrix} 1 & -1 \\ -1 & -1 \end{bmatrix} \\ x_0 &= \sqrt{2} \\ x_f &= 1 \\ \lambda_{x0} &= \frac{x_f - B_{11}x_0}{B_{12}} \end{aligned} \tag{50}$$

and

$$\mathbf{B} = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} = \exp(\mathbf{A}t_f) \tag{51}$$

Example 1 was solved using the Radau pseudospectral method (RPM) with the software *OptimalPrime*[27] and the NLP solver SNOPT[28] for $N = 4$ to $N = 99$ LGR points. The SNOPT optimality and feasibility tolerances were 10^{-10}. A typical solution for $N = 39$ LGR points (i.e., $N + 1 = 40$ discretization points) is shown in Fig. 2 alongside the exact solution. Suppose now that we define the following maximum absolute errors between the RPM solution and the exact solution:

$$\begin{aligned} e_y &= \max_{k\in[1,\ldots,N+1]} \log_{10} |y(\tau_k) - y^*(\tau_k)| \\ e_{\lambda_y} &= \max_{k\in[1,\ldots,N+1]} \log_{10} \left|\lambda_y(\tau_k) - \lambda_y^*(\tau_k)\right| \\ e_u &= \max_{k\in[1,\ldots,N]} \log_{10} |u(\tau_k) - u^*(\tau_k)| \end{aligned} \tag{52}$$

Figs. 3–4 show e_y, e_u, and e_{λ_y} as a function of $N + 1$. It is seen that e_y, e_u, and e_{λ_y} decrease in a linear manner from $N = 4$ to 49. Moreover, for $N \geq 50$ all three errors remain essentially constant, e_y and e_u being constant at approximately 10^{-10} and e_u being constant at approximately 10^{-9}, as expected from the SNOPT tolerances used. The rate of decrease of e for the lower number of nodes is most revealing because it shows that e decreases linearly, demonstrating a spectral convergence rate.

Example 2: Infinite-Horizon LQR Problem

Consider the following optimal control problem taken from Ref. 12. Denoting $\mathbf{x}(t) = [x_1(t)\ x_2(t)]^T \in \mathbb{R}^2$ as the state and $u(t) \in \mathbb{R}$ as the control, minimize the cost functional

$$J = \frac{1}{2}\int_0^\infty \left(\mathbf{x}^T\mathbf{Q}\mathbf{x} + u^T R u\right) dt, \tag{53}$$

subject to the dynamic constraint

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{B}u, \tag{54}$$

and the initial condition

$$\mathbf{x}(0) = \begin{bmatrix} -4 \\ 4 \end{bmatrix}. \tag{55}$$

The matrices $\mathbf{A}$, $\mathbf{B}$, $\mathbf{Q}$, and R for this problem are given as

$$\mathbf{A} = \begin{bmatrix} 0 & 1 \\ 2 & -1 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad \mathbf{Q} = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}, \quad R = \frac{1}{2}. \tag{56}$$

The exact solution to this problem is

$$\begin{aligned} \mathbf{x}(t) &= \exp([\mathbf{A} - \mathbf{B}\mathbf{K}]\, t)\mathbf{x}(0) \\ u(t) &= -\mathbf{K}\mathbf{x}(t) \\ \boldsymbol{\lambda}(t) &= \mathbf{S}\mathbf{x}(t) \end{aligned} \tag{57}$$

where $\mathbf{K}$ is the optimal feedback gain and $\mathbf{S}$ is the solution to the algebraic Riccati equation. In this case $\mathbf{K}$ and $\mathbf{S}$ are given, respectively, as

$$\begin{aligned} \mathbf{K} &= \begin{bmatrix} 4.828427124746193 & 2.557647291327851 \end{bmatrix} \\ \mathbf{S} &= \begin{bmatrix} 6.031273049535752 & 2.414213562373097 \\ 2.414213562373097 & 1.278823645663925 \end{bmatrix} \end{aligned} \tag{58}$$

The optimal control problem of Eqs. (53)–(55) was solved using the infinite-horizon version of the Radau pseudospectral method (as given in Section 6) using the software *OptimalPrime*[27] and the NLP solver SNOPT[28] with default optimality and feasibility tolerances of 10^{-6} and 2×10^{-6}, respectively, for $N = 4$ to $N = 34$ (i.e., $N + 1 = 5$ to $N + 1 = 35$ points) by steps of 5. The infinite-horizon RPM solution for $N+1 = 35$ is shown in Figs 6–7 as a function of τ alongside the exact solution. It is seen that the RPM solution and the exact solution are indistinguishable for all three quantities (state, control, and costate). In particular, it is seen that the infinite horizon version of the RPM solves the problem at all of the LGR points *plus* the point $\tau = +1$ (i.e., $t = \infty$), thus computing the solution on the infinite horizon. Suppose now that we define the following maximum absolute errors between the RPM solution and the exact solution:

$$\begin{aligned} e_{\mathbf{x}} &= \max_{k \in [1,\ldots,N+1]} \log_{10} \|\mathbf{x}(\tau_k) - \mathbf{x}^*(\tau_k)\| \\ e_u &= \max_{k \in [1,\ldots,N]} \log_{10} |u(\tau_k) - u^*(\tau_k)| \\ e_{\boldsymbol{\lambda}} &= \max_{k \in [1,\ldots,N+1]} \log_{10} \|\boldsymbol{\lambda}(\tau_k) - \boldsymbol{\lambda}^*(\tau_k)\| \end{aligned} \tag{59}$$

The values of $e_{\mathbf{x}}$, $e_{\boldsymbol{\lambda}_y}$, and e_u are shown in Figs. 9–11. It is seen that all errors decrease linearly until approximately $N = 35$, again demonstrating a spectral convergence rate.

8 Comparison with Previous Work on LGR Collocation

It is noted that two earlier LGR collocation methods have been presented in Refs. 12 and 16. The method of Kameswaran and Biegler in Ref. 16 focuses on *local* collocation using LGR points. The method of Fahroo and Ross in Ref. 12 describes a *global* method for solving *infinite-horizon* problems. In this section we comment briefly on how the method derived in this paper relates to this previous work.

8.1 Comparison with Local LGR Collocation Method in Ref. 16

The method derived in this paper shares similarities with the method of Ref. 16 in that the approximation of the state uses the same basis of Lagrange polynomials. It is noted, however, that the method of Ref. 16 uses *local* collocation, favoring a small number of collocation points and many subintervals (called *finite elements* in Ref. 16). The degree of the polynomials on each subinterval is fixed and convergence is achieved by increasing the number of subintervals. The current paper focuses on a global collocation method where there is a single interval and convergence is achieved by increasing the degree of the polynomials. The method of Ref. 16 leads to a sparse optimization problem with a large number of variables and constraints at the endpoints of each subinterval, while the global method in this paper leads to a low dimensional, dense optimization problem. The method of Ref. 16 is implemented similar to an implicit Runge-Kutta method (due to the fact that the time interval is divided into many subintervals) whereas the method derived in this paper is implemented in the form of a pseudospectral method. It is noted that both approaches are valid, but the current approach is consistent with the manner in which pseudospectral methods have been implemented over the past several years in the aerospace control literature.

8.2 Comparison with Global Infinite-Horizon LGR Method in Ref. 12

In the Lobatto pseudospectral approach as described in Ref. 1, the state is approximated by polynomials of degree $N-1$ and the system dynamics is collocated at the N Lobatto quadrature points. For the infinite horizon control problem studied in Ref. 12, Fahroo and Ross propose using a change of variables to map the infinite time interval onto the half-open interval $[-1,+1)$. This change of variables leads to a singularity in the transformed dynamics at $\tau=+1$. Hence, it is not possible to collocate at $\tau=+1$. To handle this singularity, Fahroo

and Ross propose to *collocate and discretize* at the Radau quadrature points for which $\tau_N < 1$. The fundamental difference between the pseudospectral scheme of Ref. 12 and the scheme introduced in this paper is that in Ref. 12, the state is approximated by polynomials of degree $N-1$, while in this paper the state is approximated using polynomials of degree N. This change in the degree of the polynomials leads to fundamental differences between the two schemes. For example, since the Lagrange polynomials are of different degrees, the differentiation matrices are completely different. The differentiation matrix used in Ref. 12 is singular, while the matrices $\mathbf{D}_{2:N+1}$ in (11) and $\mathbf{D}_{1:N}$ in (12) used in this paper are invertible. If the control and the initial state $\mathbf{x}_0$ are given, then the collocated dynamics in Ref. 12 constitutes N equations in $N-1$ unknowns $\mathbf{X}_{2:N}$, an *overdetermined* system. In contrast, (11) constitutes N equations in N unknowns $\mathbf{X}_{2:N+1}$ where the coefficient matrix $\mathbf{D}_{2:N+1}$ is invertible. In the approach of Ref. 12, $\mathbf{X}_{N+1}$, the estimate of the state at $\tau = +1$, is removed from the problem by using polynomials of degree $N-1$ instead of polynomials of degree N. In the approach presented here, the state is approximated at τ_i, $1 \leq i \leq N+1$. Hence, $\mathbf{X}_{N+1}$, the estimate of the state at the horizon, is a variable included in the pseudospectral scheme. In addition to the fact that for infinite-horizon problems the state is estimated at the horizon, the ability to estimate the state at $\tau = +1$ is useful in finite-horizon problems when the objective function depends on the state at the terminal time or when there is an endpoint constraint.

9 Conclusions

A method has been presented for direct trajectory optimization and costate estimation using global collocation at Legendre-Gauss-Radau (LGR) points. A theoretical foundation for the method has been provided. The method can be viewed either as a global implicit integration method or a pseudospectral method. Using the pseudospectral (or differential) form, it is possible to solve general optimal control problems and construct a complete mapping between the continuous and discrete variables. The method presented in this paper has been demonstrated on both a finite-horizon and infinite-horizon control problem, thereby demonstrating the range of its utility. The results of this paper indicate that the Radau pseudospectral method described in this paper leads to the ability to determine accurate primal and dual solutions to general optimal control problems.

Acknowledgments

The authors gratefully acknowledge support for this research from the U.S. Army Research Office under Grant 55173-CI, The Office of Naval Research under Grant N00014-08-1-1173, and from the National Science Foundation under Grants 0619080 and 0620286.

References

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

[2] Elnagar, G. and Kazemi, M., "Pseudospectral Chebyshev Optimal Control of Constrained Nonlinear Dynamical Systems," *Computational Optimization and Applications*, Vol. 11, No. 2, 1998, pp. 195–217.

[3] Fahroo, F. and Ross, I. M., "A Spectral Patching Method for Direct Trajectory Optimization," *Journal of Astronautical Sciences*, Vol. 48, No. 2-3, Apr.-Sept. 2000, pp. 269–286.

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

[5] Ross, I. M. and Fahroo, F., *Lecture Notes in Control and Information Sciences*, chap. Legendre Pseudospectral Approximations of Optimal Control Problems, Springer-Verlag, 2003.

[6] Rao, A. V., "Extension of a Pseudospectral Legendre Method for Solving Non-Sequential Multiple-Phase Optimal Control Problems," AIAA Guidance, Navigation, and Control Conference, *AIAA Paper 2003-5634*, Austin, Texas, August 11–14 2003.

[7] Williams, P., "Jacobi Pseudospectral Method for Solving Optimal Control Problems," *Journal of Guidance, Control, and Dynamics*, Vol. 27, No. 2, 2004, pp. 293–297.

[8] Williams, P., "Application of Pseudospectral methods for Receding Horizon Control," *Journal of Guidance, Control, and Dynamics*, Vol. 27, No. 2, 2004, pp. 310–314.

[9] Williams, P., "Hermite-Legendre-Gauss-Lobatto Direct Transcription Methods in Trajectory Optimization," *Americal Astronautical Society*, Spaceflight Mechanics Meeting, August 2005.

[10] Ross, I. M. and Fahroo, F., "Pseudospectral Knotting Methods for Solving Optimal Control Problems," *Journal of Guidance, Control, and Dynamics*, Vol. 27, No. 3, 2004, pp. 397–405.

[11] Fahroo, F. and Ross, I. M., "On Discrete-Time Optimality Conditions for Pseudospectral Methods," *AIAA Guidance, Navigation, and Control Conference*, AIAA Paper 2006-6304, Keystone, Colorado, August 2006.

[12] Fahroo, F. and Ross, I. M., "Pseudospectral Methods for Infinite-Horizon Nonlinear Optimal Control Problems," *Journal of Guidance, Control, and Dynamics*, Vol. 31, No. 4, 2008, pp. 927–936.

[13] Benson, D. A., *A Gauss Pseudospectral Transcription for Optimal Control*, Ph.D. thesis, MIT, 2004.

[14] Benson, D. A., Huntington, G. T., Thorvaldsen, T. P., and Rao, A. V., "Direct Trajectory Optimization and Costate Estimation via an Orthogonal Collocation Method," *Journal of Guidance, Control, and Dynamics*, Vol. 29, No. 6, November-December 2006, pp. 1435–1440.

[15] Huntington, G. T., *Advancement and Analysis of a Gauss Pseudospectral Transcription for Optimal Control*, Ph.D. thesis, Department of Aeronautics and Astronautics, Massachusetts Institute of Technology, 2007.

[16] Kameswaran, S. and Biegler, L. T., "Convergence Rates for Direct Transcription of Optimal Control Problems Using Collocation at Radau Points," *Computational Optimization and Applications*, Vol. 41, No. 1, 2008, pp. 81–126.

[17] Reddien, G. W., "Collocation at Gauss Points as a Discretization in Optimal Control," *SIAM Journal on Control and Optimization*, Vol. 17, No. 2, March 1979.

[18] Cuthrell, J. E. and Biegler, L. T., "On the Optimization of Differential-Algebraic Processes," *AIChe Journal*, Vol. 33, No. 8, August 1987, pp. 1257–1270.

[19] Cuthrell, J. E. and Biegler, L. T., "Simultaneous Optimization and Solution Methods for Batch Reactor Control Profiles," *Computers and Chemical Engineering*, Vol. 13, No. 1/2, 1989, pp. 49–62.

[20] Ross, I. M. and Fahroo, F., "Advances in Pseusospectral Methods for Optimal Control," *AIAA Guidance, Navigation, and Control Conference*, AIAA Paper 2008-7309, Honolulu, Hawaii, August 2008.

[21] Ross, I. M. and Fahroo, F., "Convergence of the Costates Do Not Imply Convergence of the Controls," *Journal of Guidance, Control, and Dynamics*, Vol. 31, No. 4, 2008, pp. 1492–1497.

[22] Huntington, G. T., Benson, D. A., and Rao, A. V., "Optimal Configuration of Tetrahedral Spacecraft Formations," *The Journal of the Astronautical Sciences*, Vol. 55, No. 2, April-June 2007, pp. 141–169.

[23] Huntington, G. T. and Rao, A. V., "Optimal Reconfiguration of Spacecraft Formations Using the Gauss Pseudospectral Method," *Journal of Guidance, Control, and Dynamics*, Vol. 31, No. 3, May-June 2008, pp. 689–698.

[24] Hager, W. W., "Numerical analysis in optimal control," *International Series of Numerical Mathematics*, edited by K.-H. Hoffmann, I. Lasiecka, G. Leugering, J. Sprekels, and F. Troeltzsch, Vol. 139, Birkhauser Verlag, Basel/Switzerland, 2001, pp. 83–93.

[25] Hager, W. W., Dontchev, A., Poore, A., and Yang, B., "Optimality, stability, and convergence in nonlinear control," *Applied Math. and Optim.*, Vol. 31, 1995, pp. 297–326.

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

[27] Patterson, M. A., "*OptimalPrime:* A MATLAB Software for Solving Non-Sequential Multiple-Phase Optimal Control Problems Using Pseudospectral Methods," Dept. of Mechanical and Aerospace Engineering, University of Florida, August 2008.

[28] Gill, P. E., Murray, W., and Saunders, M. A., *User's Guide for SNOPT Version 7: Software for Large Scale Nonlinear Programming*, February 2006.

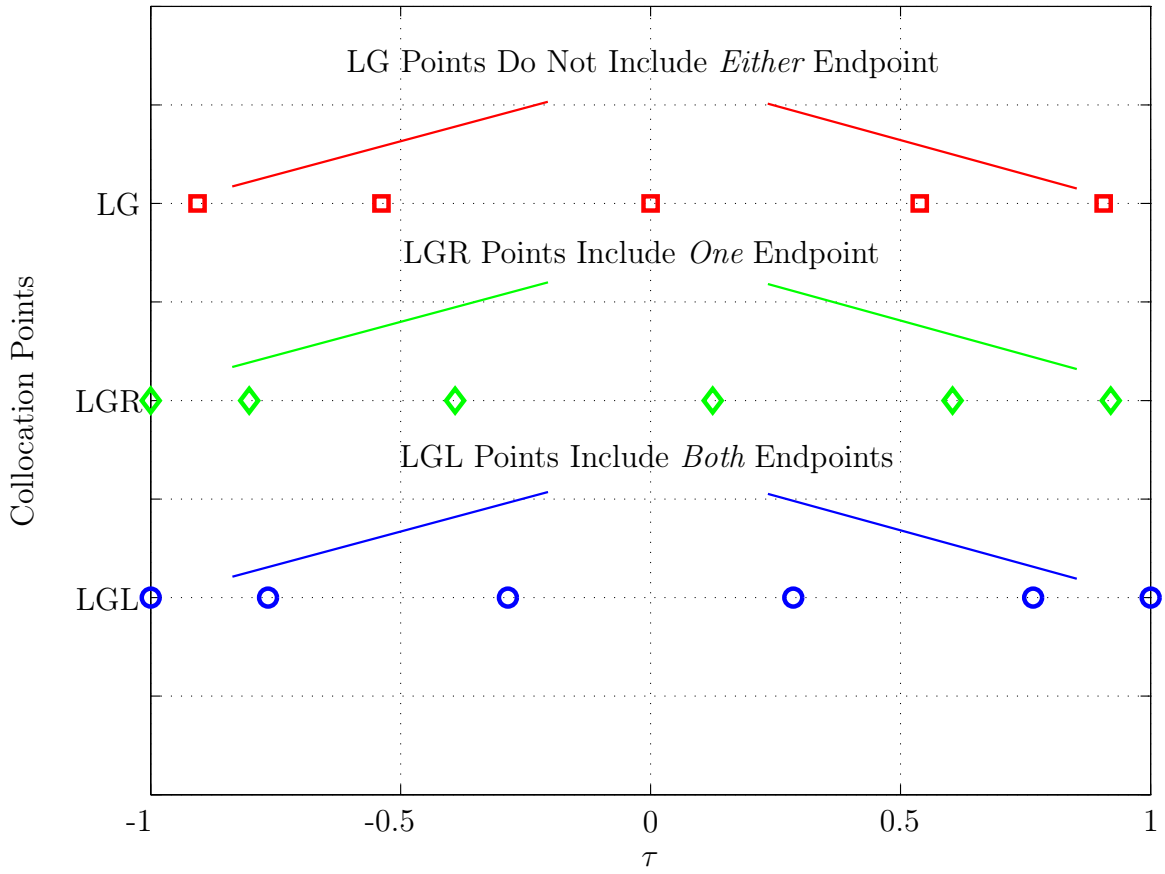

Figure 1: Schematic Showing the Differences Between LGL, LGR, and LG Collocation Points.

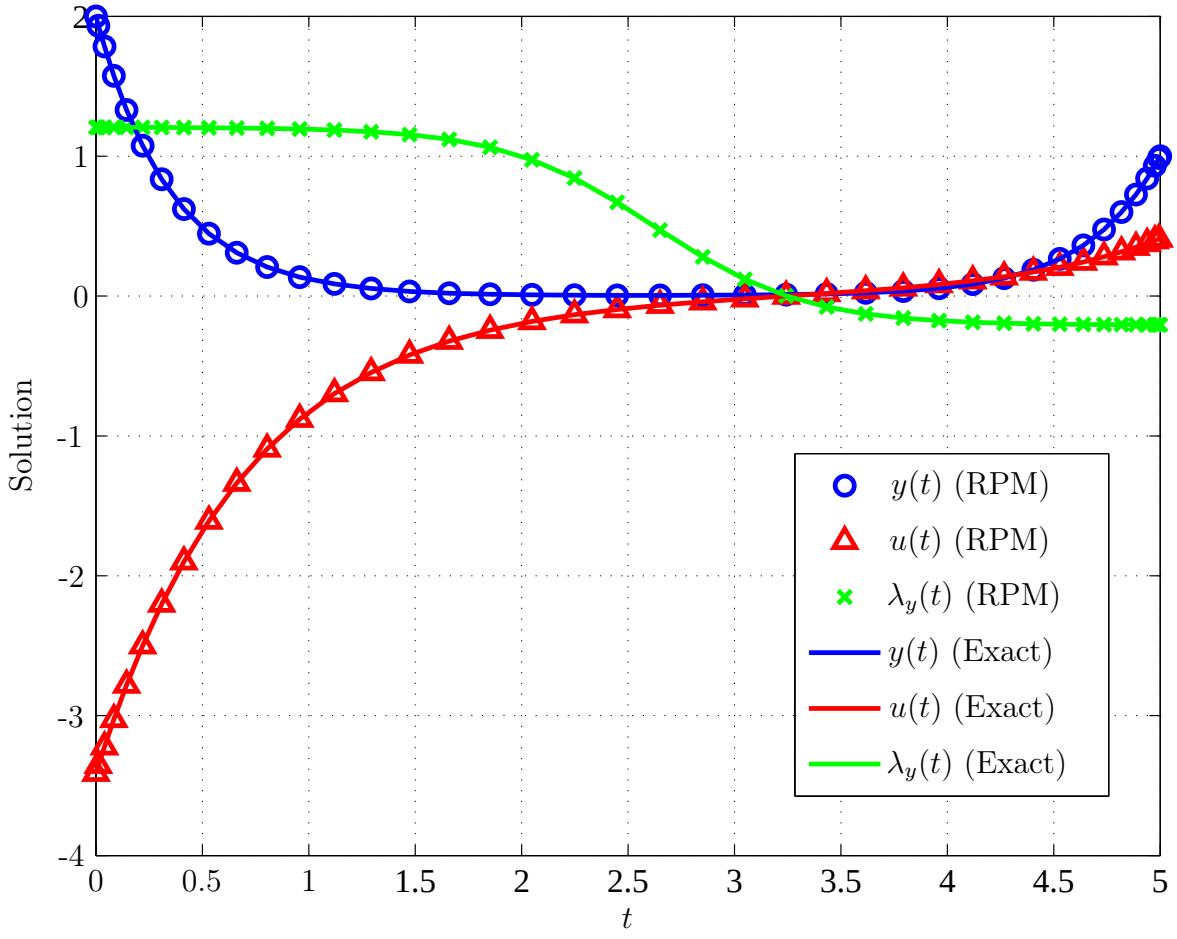

Figure 2: Solution to Example 1 Using 39 LGR Points Alongside Exact Solution.

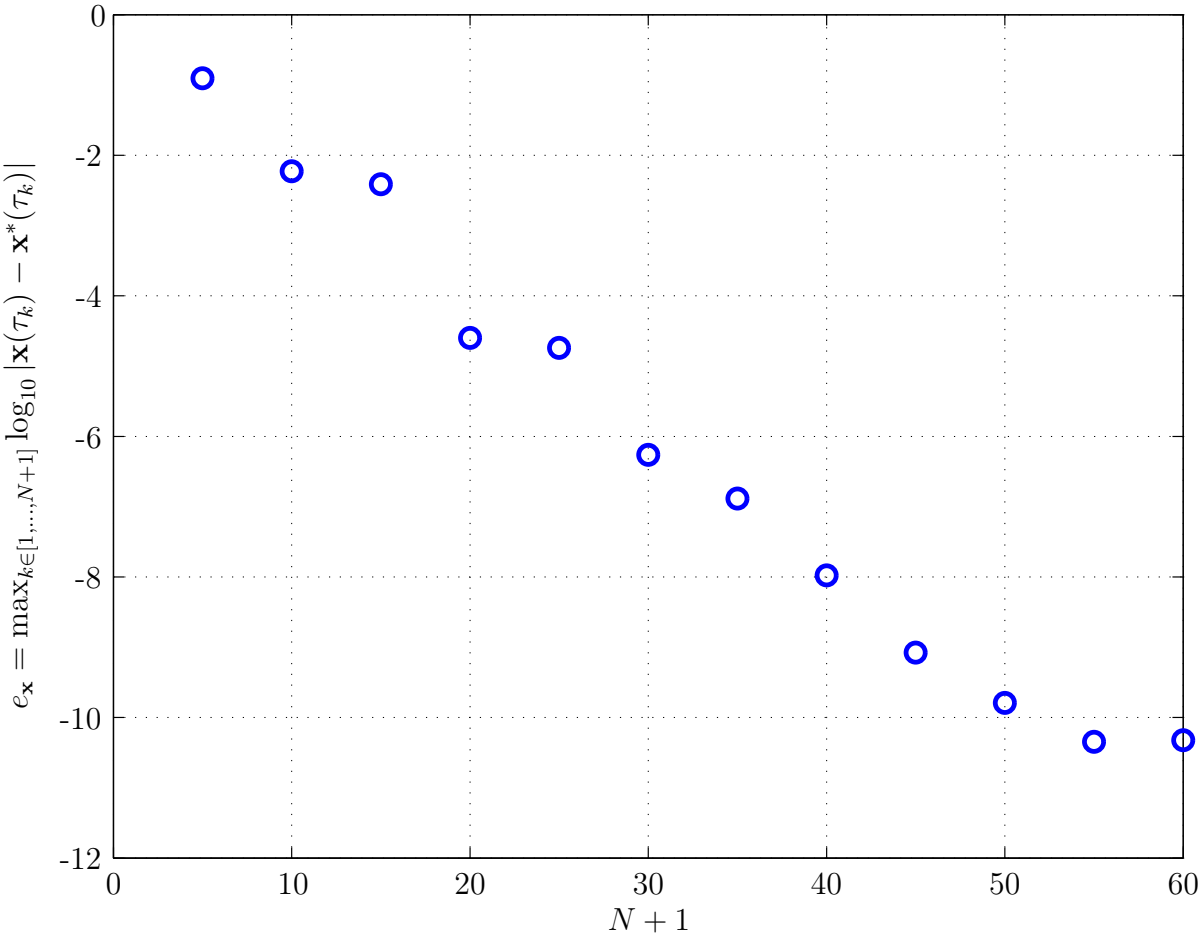

Figure 3: Error in Radau Pseudospectral State for Example 1.

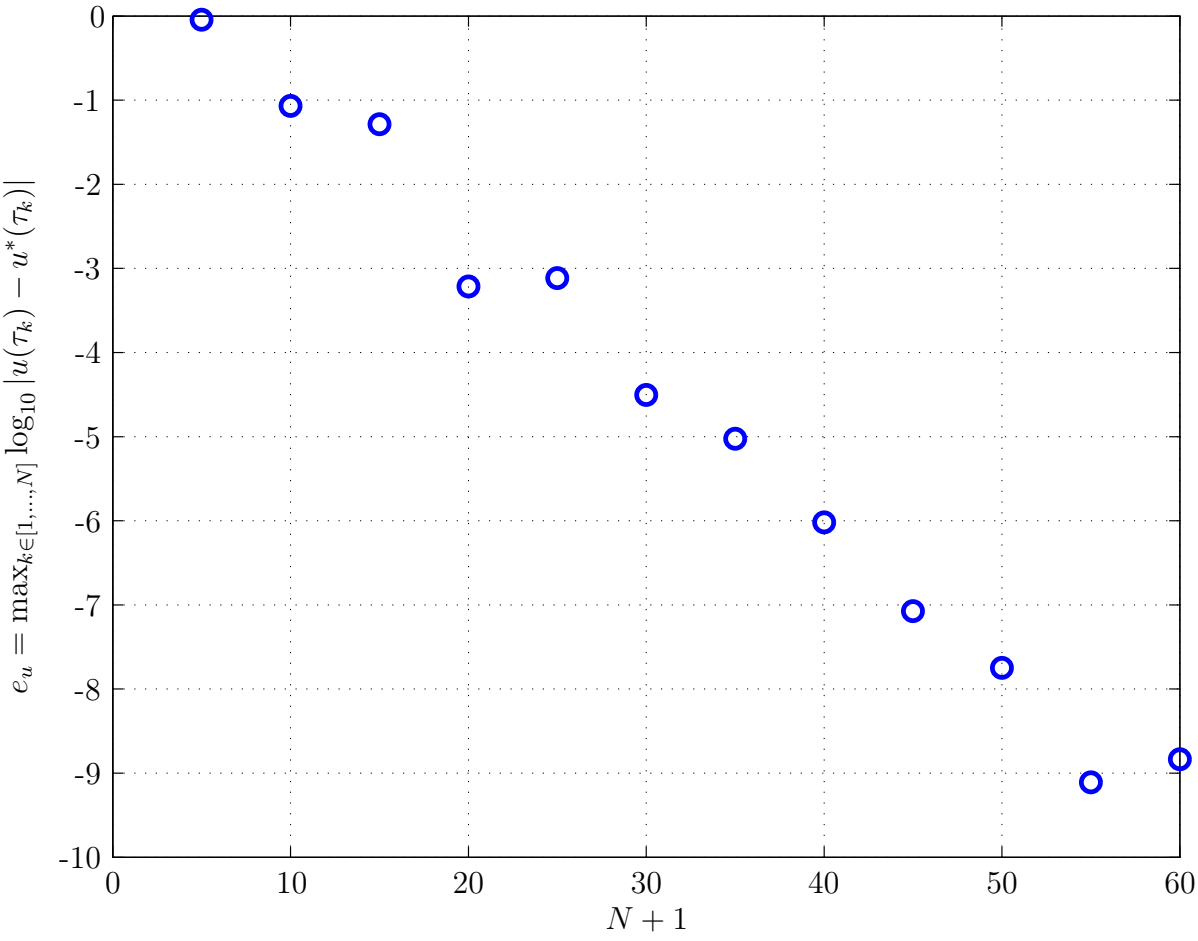

Figure 4: Error in Radau Pseudospectral Control for Example 1.

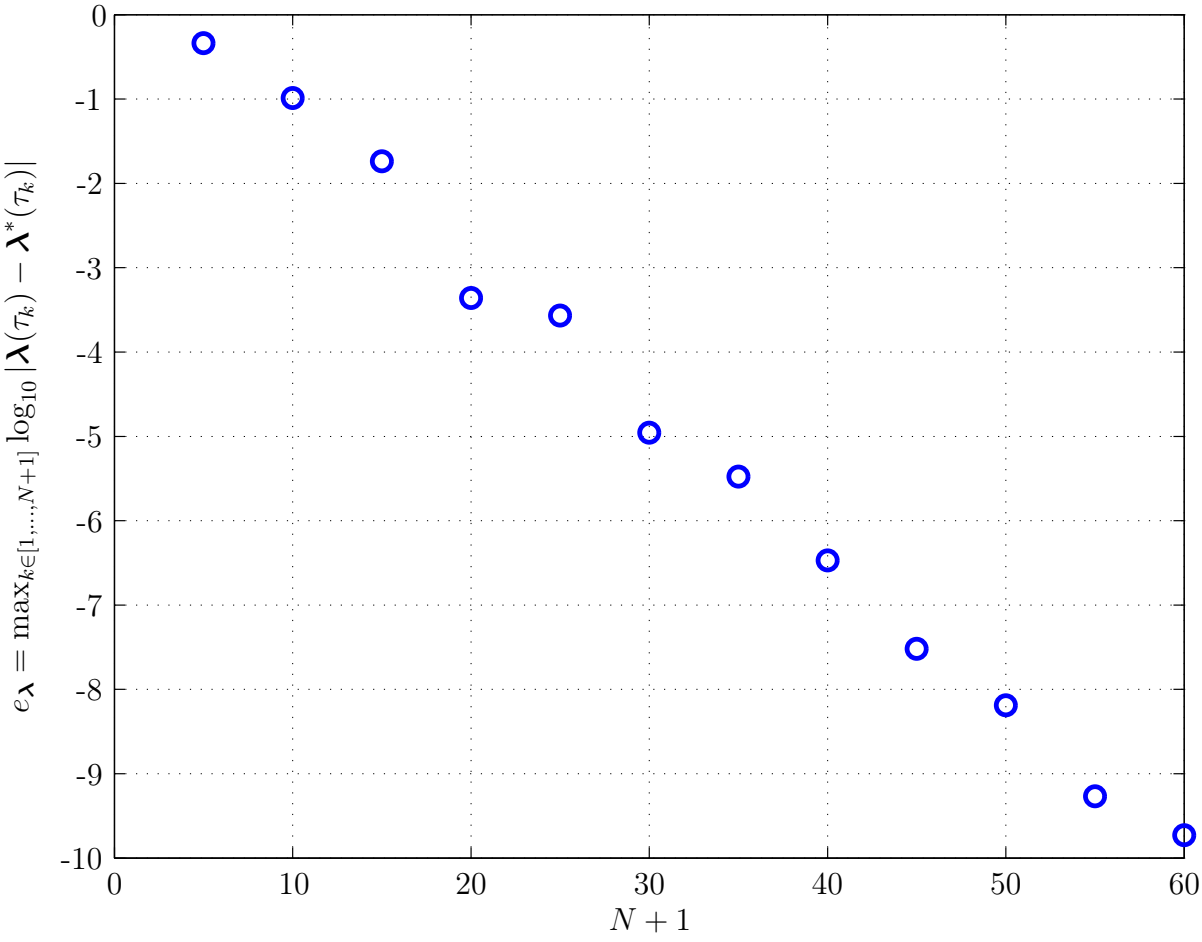

Figure 5: Error in Radau Pseudospectral Costate for Example 1.

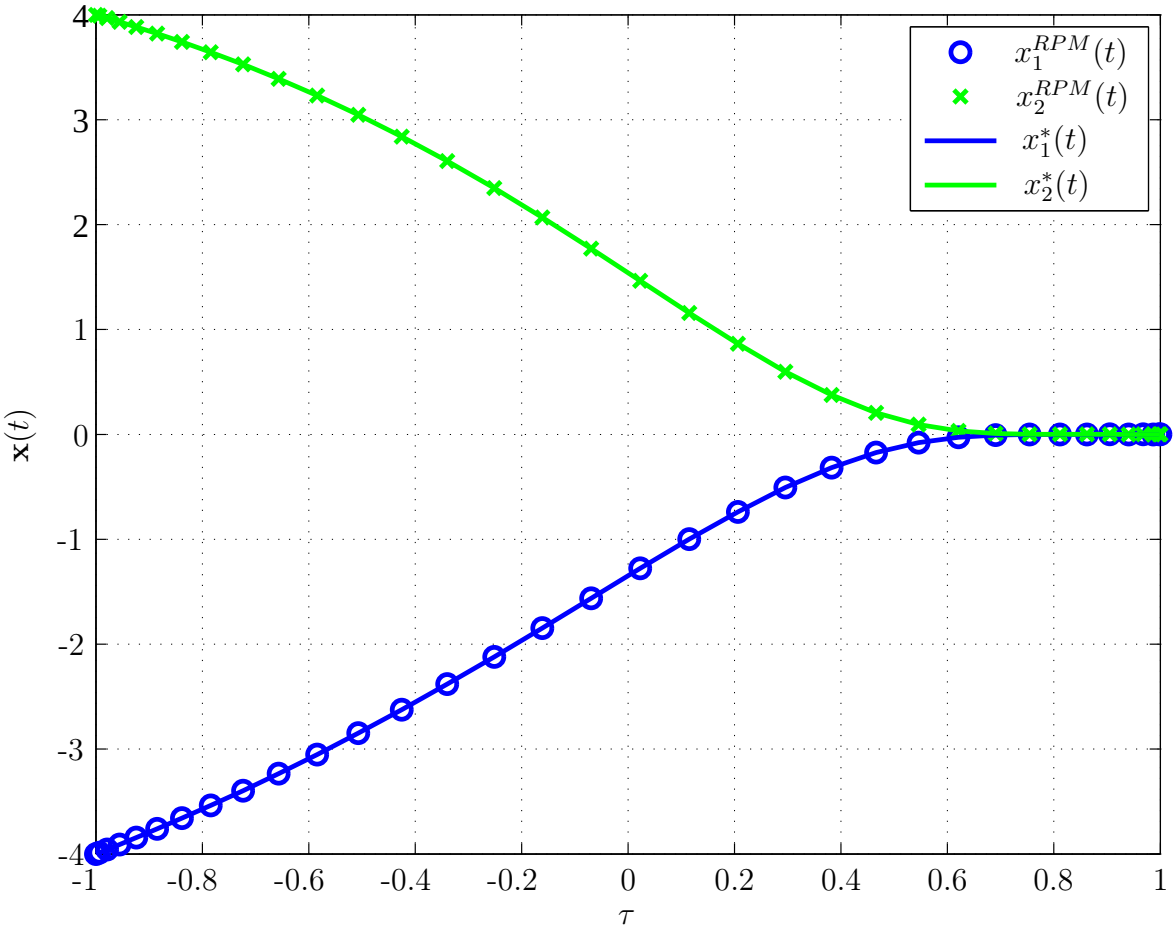

Figure 6: Infinite-Horizon Radau Pseudospectral State Solution for Example 2 Using $N = 34$ $(N + 1 = 35)$ as a Function of τ Alongside Exact Solution.

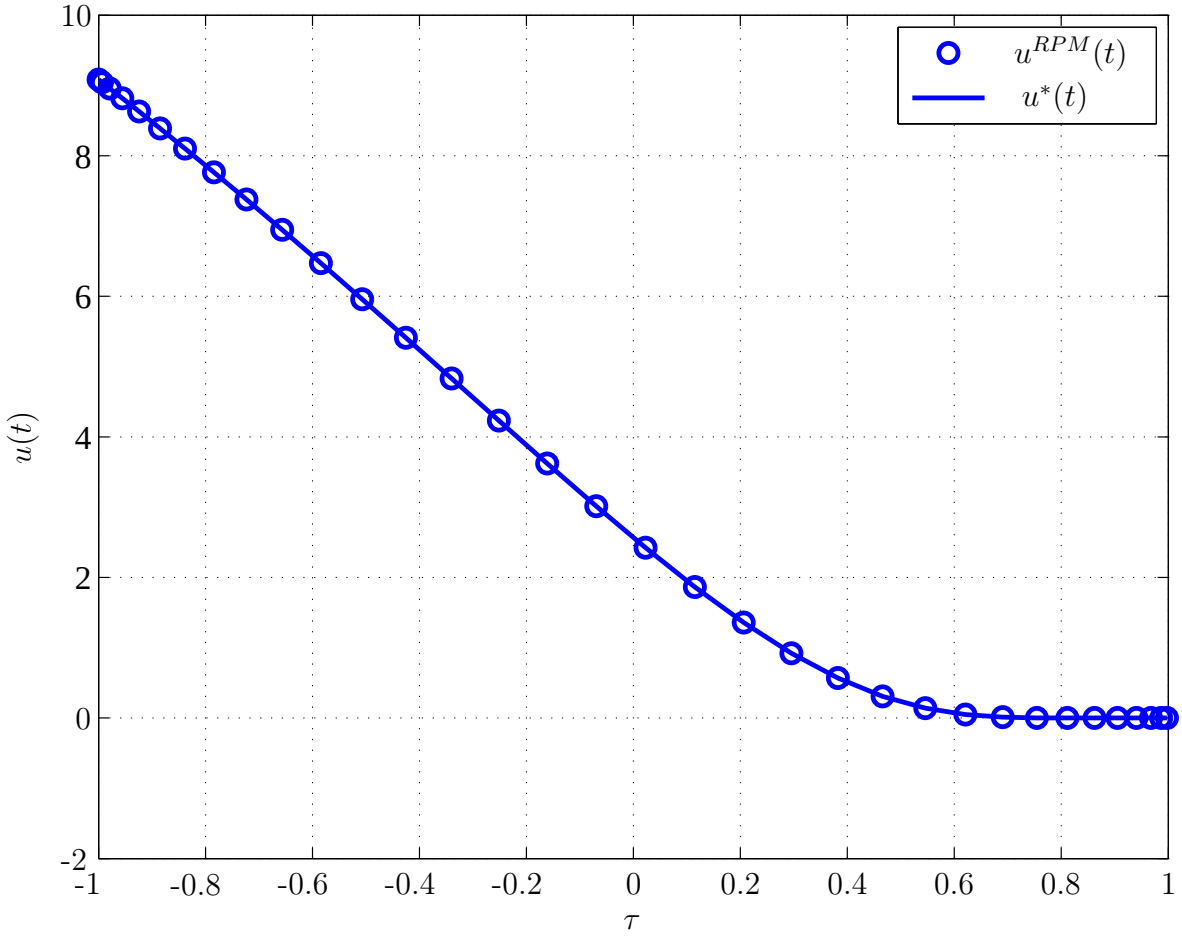

Figure 7: Infinite-Horizon Radau Pseudospectral Control Solution for Example 2 Using $N = 39$ $(N + 1 = 40)$ as a Function of τ Alongside Exact Solution.

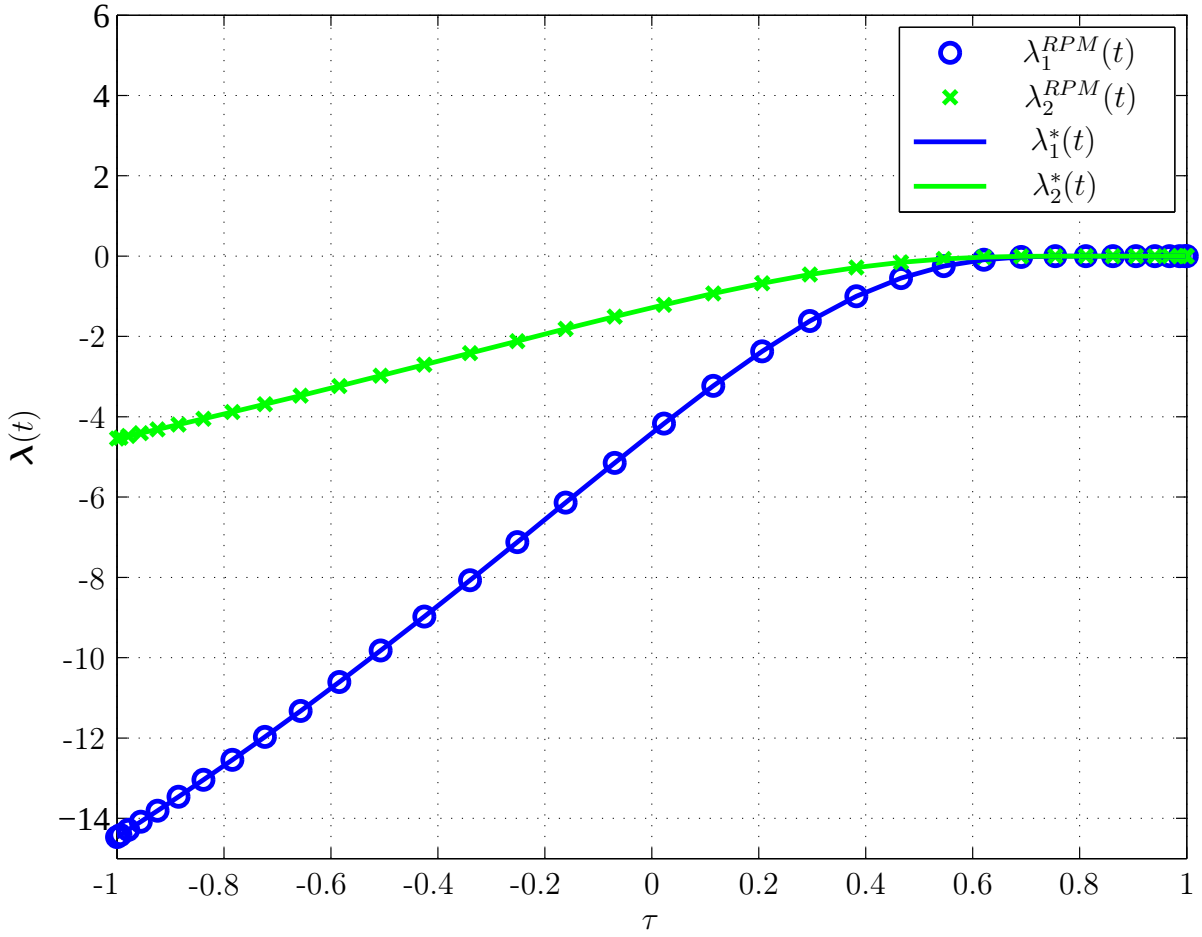

Figure 8: Infinite-Horizon Radau Pseudospectral Costate Solution for Example 2 Using $N = 39$ $(N + 1 = 40)$ as a Function of τ Alongside Exact Solution.

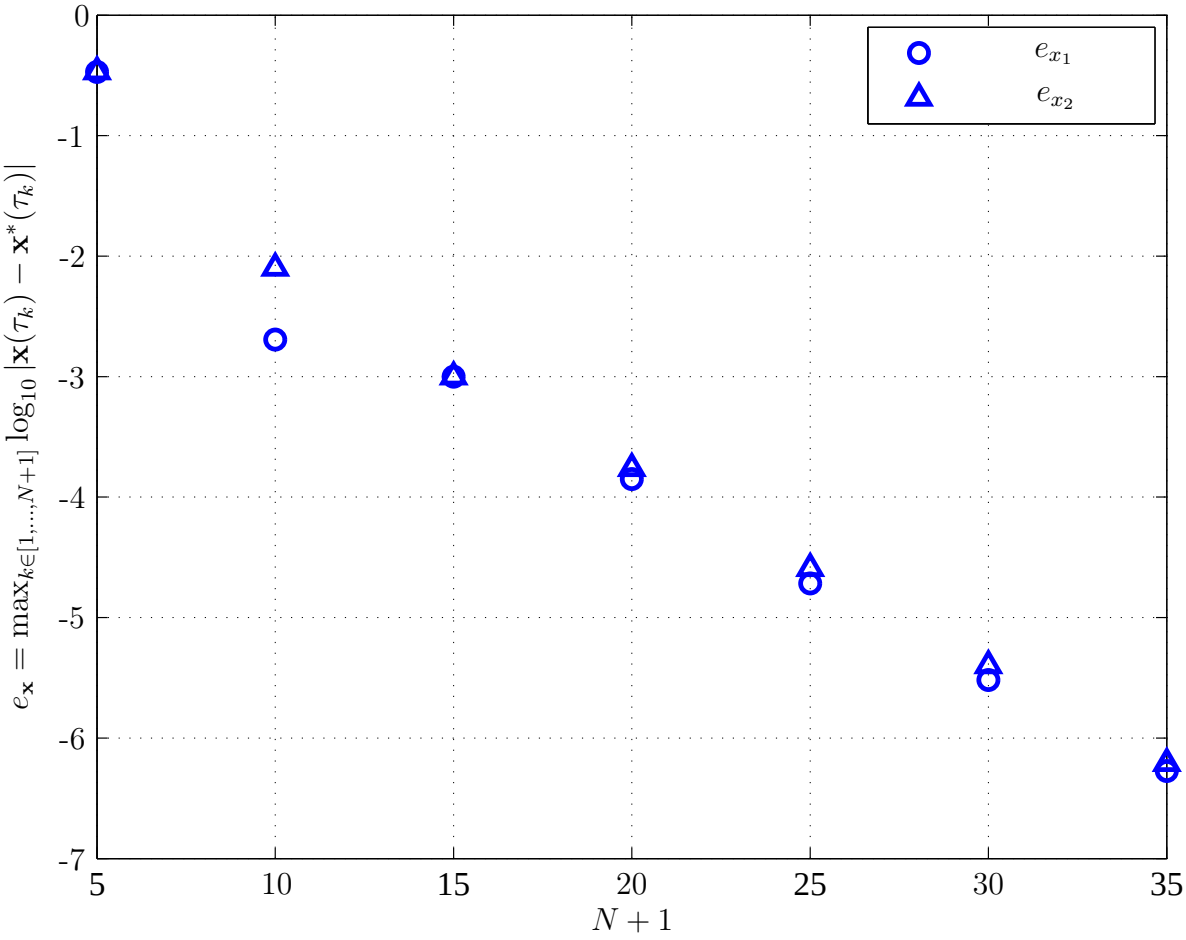

Figure 9: Error in Infinite-Horizon Radau Pseudospectral State Solution as a function of $N+1$.

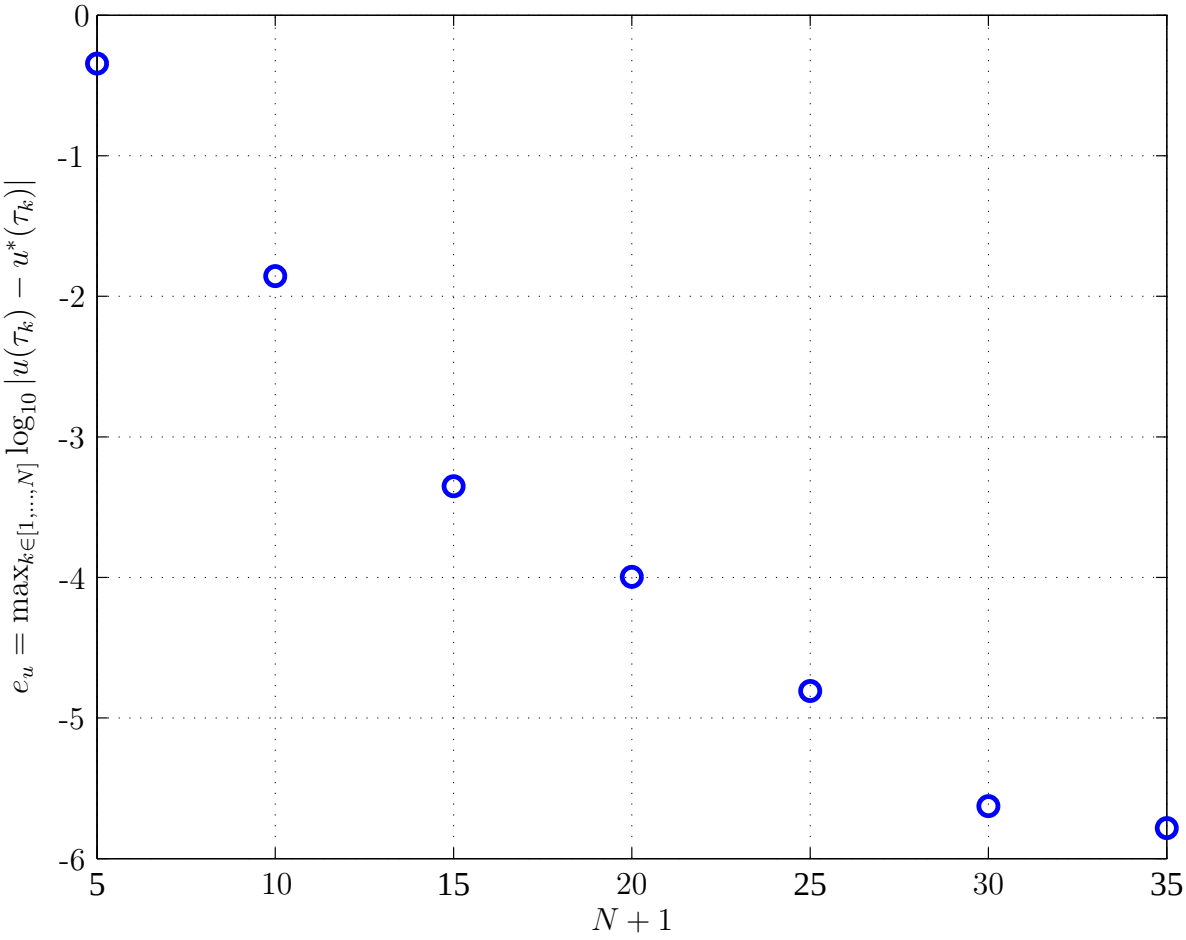

Figure 10: Error in Infinite-Horizon Radau Pseudospectral Control Solution as a function of $N+1$.

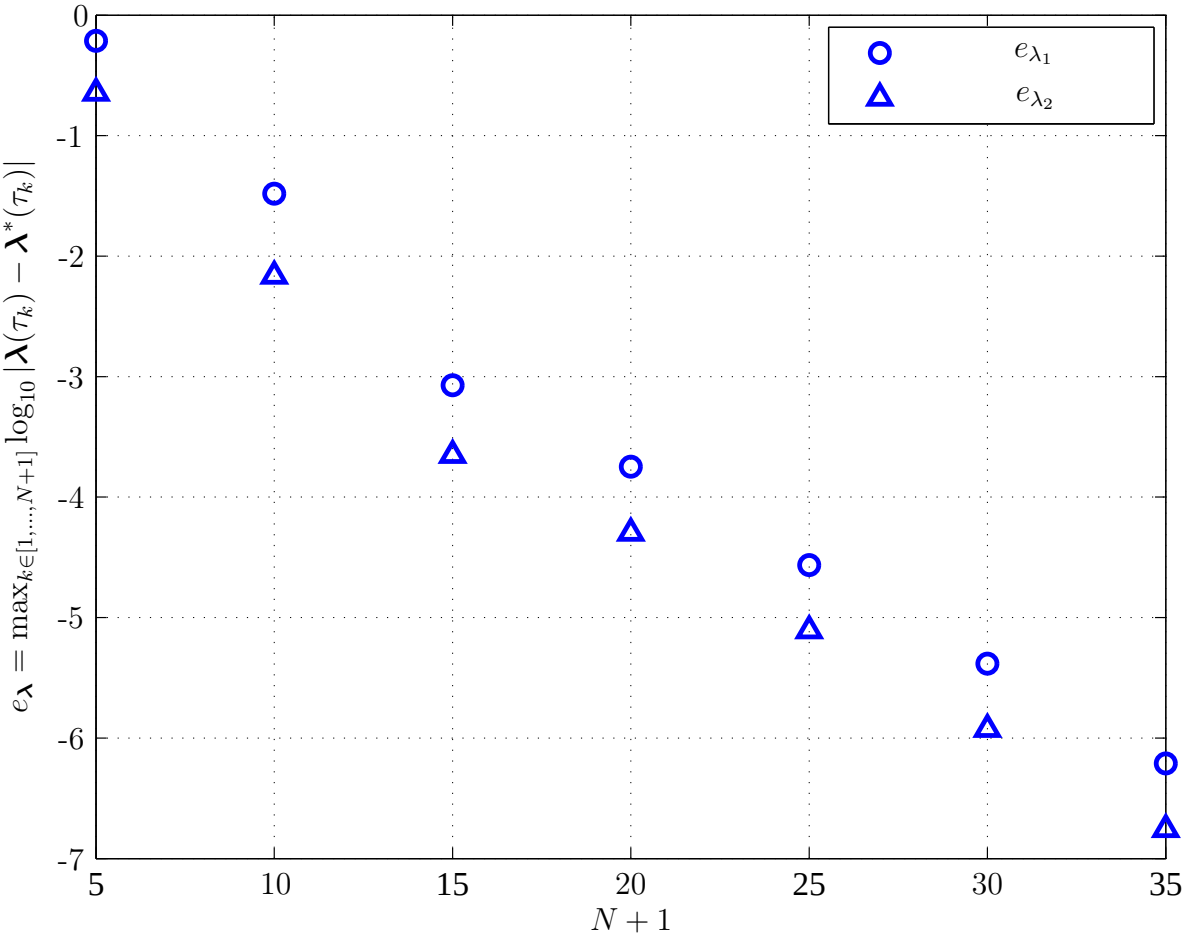

Figure 11: Error in Infinite-Horizon Radau Pseudospectral Costate Solution as a function of $N+1$.