A Comparison of Accuracy and Computational Efficiency of Three Pseudospectral Methods

Geoffrey T. Huntington*
David Benson†
Charles Stark Draper Laboratory
555 Technology Square
Cambridge, MA 02139

Anil V. Rao‡
University of Florida
Gainesville, FL 32611-6250

Abstract

A comparison is made between three pseudospectral methods used to numerically solve optimal control problems. In particular, the accuracy of the state, control, and costate obtained using the Legendre, Radau, and Gauss pseudospectral methods is compared. Three examples with different degrees of complexity highlight the differences in performance between the three methods. The results of this study indicate that all three methods are comparable in state accuracy, yet differ in their costate accuracy. These differences are analytically supported by rigorous mathematical analysis. The Gauss and Radau pseudospectral methods are shown to produce significantly more accurate costate estimations than the Legendre method at no additional computational burden. Based on these results, a rationale is created to determine when it is appropriate to use each method to solve optimal control problems.

1 Introduction

Numerical methods for solving optimal control problems fall into two general categories: indirect methods and direct methods. In an indirect method, the first-order optimality conditions are derived using the calculus of variations[1] and Pontryagin's minimum principle.[2] These optimality conditions lead to a Hamiltonian boundary-value problem (HBVP) which is then solved numerically to determine extremal trajectories. In a direct method, the original optimal control problem is transcribed to a nonlinear programming problem (NLP) which is solved numerically using a well-established optimization method. In recent years, direct methods for solving optimal control problems have become increasingly popular over indirect methods for several reasons. First, when using a direct method it is unnecessary to derive the (generally quite complex) first-order optimality conditions (i.e. HBVP). Second, even in the case where deriving the HBVP is straightforward, solving the HBVP can be extremely difficult due to the high sensitivity of the Hamiltonian dynamics to unknown boundary conditions. Contrariwise, direct methods require much less mathematical derivation and the resulting NLP can be solved with much greater ease as compared to the HBVP. For these reasons, direct methods are often the preferred approach for numerically solving complex optimal control problems.[3]

The category of direct methods is quite broad and encompasses some very different techniques.[4] In particular, the choice of what quantities to discretize and how to approximate the continuous-time dynamics varies widely amongst the different direct methods. Two of the more common direct methods are *control parameterization* and *state and control parameterization*[5] methods. In a control parameterization method, the control alone is approximated and the differential equations are solved via numerical integration. Examples

*Ph.D., Dept. of Aeronautics and Astronautics, MIT. E-mail: ghuntington@alum.mit.edu
†Senior Member of the Technical Staff, GN&C Systems Division. E-mail: dbenson@draper.com
‡Assistant Professor, Department of Mechanical and Aerospace Engineering, University of Florida, P.O. Box 116250, Gainesville, FL, 32611. E-mail: anilvrao@ufl.edu. Corresponding Author.

of control parameterization include shooting methods and multiple shooting methods. In state and control parameterization methods, the state is discretized in the NLP as well as the control, and the continuous-time differential equations are converted into algebraic constraints. These constraints are then imposed in the NLP formulation, which avoid the sensitivity issues of direct shooting methods at the expense of a larger NLP.

In recent years, considerable attention has been focused on a class of state and control parameterization methods called *pseudospectral*[6–8] or *orthogonal collocation*[9,10] methods. In a pseudospectral method, a finite basis of global interpolating polynomials is used to approximate the state and control at a set of discretization points. The time derivative of the state in the dynamic equations is approximated by the derivative of the interpolating polynomial and is then constrained to be equal to the vector field of the dynamic equations at a set of collocation points. While any set of unique collocation points can be chosen, generally speaking an *orthogonal collocation* is chosen, i.e. the collocation points are chosen to be the roots of an orthogonal polynomial (or linear combinations of such polynomials and their derivatives). Because pseudospectral methods are generally implemented via orthogonal collocation, the terms *pseudospectral* and *orthogonal collocation* are essentially interchangeable (thus researchers in one field use the term *pseudospectral*[8] while others use the term *orthogonal collocation*[9]). One advantage to pseudospectral methods is that for smooth problems, pseudospectral methods typically have faster convergence rates than other methods, exhibiting so called "spectral accuracy".[11] For non-smooth problems or problems where modeling changes are desired, the optimal control problem can be divided into phases and orthogonal collocation can be applied globally within each phase. A vast amount of work has been done on using pseudospectral methods to solve non-smooth optimal control problems (see Refs. 10 and Refs. 12–17).

Pseudospectral methods in optimal control arose from spectral methods which were traditionally used to solve fluid dynamics problems.[7,8] Meanwhile, seminal work in orthogonal collocation methods in optimal control date back to 1979 with the work of Ref. 18 and some of the first work using orthogonal collocation methods in engineering can be found in the chemical engineering literature.[12] More recent work in chemical and aerospace engineering have used collocation at the Legendre-Gauss-Radau (LGR) points.[19–21] Because the LGR collocation technique has many similarities to other pseudospectral methods, in this paper we adopt the term *Radau pseudospectral method* (RPM) for LGR collocation. Within the aerospace engineering community, several well-known pseudospectral methods have been developed for solving optimal control problems such as the *Chebyshev pseudospectral method* (CPM),[22,23] the *Legendre pseudospectral method* (LPM),[6] and the *Gauss pseudospectral method* (GPM).[24,38,39] The CPM uses Chebyshev polynomials to approximate the state and control, and performs orthogonal collocation at the Chebyshev-Gauss-Lobatto (CGL) points. An enhancement to the Chebyshev pseudospectral method that uses a Clenshaw-Curtis quadrature was developed in Ref. 30. The LPM uses Lagrange polynomials for the approximations, and Legendre-Gauss-Lobatto (LGL) points for the orthogonal collocation. A costate estimation procedure for the Legendre pseudospectral method was developed by in Ref. 25 and recently updated in Ref. 26. Recent work by Williams shows several variants of the standard LPM. The Jacobi pseudospectral method[27] is a more general pseudospectral approach that uses Jacobi polynomials to find the collocation points, of which Legendre polynomials are a subset. Another variant, called the Hermite-LGL method,[28] uses piecewise cubic polynomials rather than Lagrange polynomials, and collocates at a subset of the LGL points.

Two key features of pseudospectral methods are the sets of *discretization* points and *collocation* points. Discretization points (or *nodes*) are the points used to discretize the state that are fed into the nonlinear program. Collocation points are the points used to discretize the control and to collocate the differential equations to ensure the dynamics have been met. These two sets of points affect the performance of the methods and the mathematical structure of the problem. For example, the specific sets of points in the GPM results in a mathematical equivalence between the Karush-Kuhn-Tucker (KKT) conditions of the NLP and the discretized optimality conditions for the HBVP at the LG points. In other words, the solution of the NLP (direct approach) also satisfies the same first-order conditions if one were using an indirect approach. Although an exact equivalence has not been proven for the other methods, much research has been done on both the LPM and the RPM that discusses the (rapid) convergence rates of the NLP solution to the solution of the continuous-time optimality conditions.[19,26,31] Moreover, an approximate relationship between the KKT conditions and the discretized first-order necessary conditions has been documented in Ref. 19 for the RPM, and in Ref. 25 and 26 for the LPM. The accuracy of these costate approximation procedures, as well as the state and control approximations, are compared in this paper.

Numerical comparisons between methods are often difficult to perform fairly because many direct methods are integrated with a specific NLP solver, thereby making it difficult to distinguish between differences between the methods and differences due to the NLP solvers. This is perhaps why there are relatively few

papers that actually compare the performance of several direct methods. In fact, even survey papers, such as that found in Ref. 3, stop short of providing comparisons between various methods. However, some research has been done that compares direct methods. Ref. 32 is primarily focused on presenting enhanced mesh refinement and scaling strategies, but contains a brief comparison of the accuracy of direct implicit integration schemes (like Hermite-Simpson integration) and pseudospectral methods. Williams created the software tool DIRECT in order to more effectively compare various trajectory optimization algorithms. The results of a thorough comparison involving 21 test problems of several different direct methods is presented in Ref. 29. Ref. 33 presents a comparison between several pseudospectral approaches, but uses unorthodox variants of pseudospectral methods that are inconsistent with a large majority of the literature.

In this paper a comparison is made between the Legendre, Radau, and Gauss pseudospectral methods. In order to provide a fair comparison, the NLP solver SNOPT[40] is used for each of the discretizations. Furthermore, all three methods are implemented using the same version of MATLAB® and the initial guesses provided for all examples are equivalent. The goal of the study is to assess the similarities and differences in the accuracy and computational performance between the three methods. Three examples are presented to make the comparison. The first two examples are designed to be sufficiently simple so that the key features of each method can be identified. The third example, a well-researched problem in aerospace engineering, is designed to provide an assessment as to how the three methods compare on a more realistic problem. A great deal of the motivation for this work is to understand when one method may perform better than another method and to identify why such a difference in performance is exhibited in such circumstances.

The results of this study indicate that the Radau and Gauss pseudospectral methods are highly comparable in numerical accuracy, computational efficiency, and ease of use. This study also shows that, even on simple problems, both the Radau and Gauss methods produce higher accuracy costate solutions as compared to the Legendre method when using the established method of Ref. 25. The authors recognize that improvements to these methods are being made, therefore an additional computation is performed using a recent advancement of the Legendre pseudospectral method. The augmented set of KKT conditions[26] is solved using the Legendre method with the covector mapping theorem, and the resulting costate is compared with the Radau and the Gauss methods. While it is found that this new costate is significantly more accurate than the original costate obtained using the method of Ref. 25, the approach is highly intractable because it requires intimate knowledge of optimal control theory and significantly increases the computation burden placed on the NLP. The complexity involved with formulating this primal-dual problem essentially defeats one of the primary advantages of using a direct method, especially when it is unnecessary in the other two approaches presented here. Lastly, analysis of the mathematical structure within each method suggests why these methods perform as they do, and suggestions are made regarding the circumstances under which each method is appropriate to use.

2 Continuous Mayer Problem

Without loss of generality, consider the following optimal control problem in Mayer form. Determine the control, $\mathbf{u}(\tau) \in \mathbb{R}^m$, that minimizes the cost functional

$$J = \Phi(\mathbf{x}(\tau_f)) \tag{1}$$

subject to the constraints

$$\frac{d\mathbf{x}}{d\tau} = \frac{t_f - t_0}{2}\mathbf{f}(\mathbf{x}(\tau), \mathbf{u}(\tau)) \in \mathbb{R}^n \tag{2}$$

$$\boldsymbol{\phi}(\mathbf{x}(\tau_0), \mathbf{x}(\tau_f)) = \mathbf{0} \in \mathbb{R}^q \tag{3}$$

The optimal control problem of Eqs. (1)–(3) is referred to as the *continuous Mayer problem*. It is noted that, in order to simplify the analysis and provide a clear comparison, this work considers optimal control problems with fixed terminal times and *without* algebraic path constraints. Furthermore, it is noted that the time interval $\tau \in [-1, 1]$ can be transformed to the the time interval $t \in [t_0, t_f]$ via the affine transformation

$$t = \frac{t_f - t_0}{2}\tau + \frac{t_f + t_0}{2} \tag{4}$$

3 First-Order Necessary Conditions of Continuous Mayer Problem

The indirect approach to solving the continuous Mayer problem of Eqs. (1)–(3) in Section 2 is to apply the calculus of variations to obtain first-order necessary conditions for optimality.[1] These variational conditions

are derived using the Hamiltonian, $\mathcal{H}$, defined for the Mayer problem as

$$\mathcal{H}(\mathbf{x}, \boldsymbol{\lambda}, \mathbf{u}) = \boldsymbol{\lambda}^T \mathbf{f}(\mathbf{x}, \mathbf{u}) \tag{5}$$

where $\boldsymbol{\lambda}(\tau) \in \mathbb{R}^n$ is the costate. Note that, for brevity, the explicit dependence of the state, control, and costate on τ has been dropped. Assuming that the optimal control lies within the feasible control set, the continuous-time first-order optimality conditions are

$$\begin{aligned}
\frac{d\mathbf{x}}{d\tau}^T &= \frac{t_f - t_0}{2}\mathbf{f}^T(\mathbf{x}, \mathbf{u}) &&= \frac{t_f - t_0}{2}\frac{\partial \mathcal{H}}{\partial \boldsymbol{\lambda}} \\
\frac{d\boldsymbol{\lambda}}{d\tau}^T &= -\frac{t_f - t_0}{2}\boldsymbol{\lambda}^T \frac{\partial \mathbf{f}}{\partial \mathbf{x}} &&= -\frac{t_f - t_0}{2}\frac{\partial \mathcal{H}}{\partial \mathbf{x}} \\
\mathbf{0}^T &= \boldsymbol{\lambda}^T \frac{\partial \mathbf{f}}{\partial \mathbf{u}} &&= \frac{\partial \mathcal{H}}{\partial \mathbf{u}} \\
& \boldsymbol{\phi}(\mathbf{x}(\tau_0), \mathbf{x}(\tau_f)) &&= \mathbf{0} \\
\boldsymbol{\lambda}(\tau_0)^T &= \boldsymbol{\nu}^T \frac{\partial \boldsymbol{\phi}}{\partial \mathbf{x}(\tau_0)} \quad , \quad \boldsymbol{\lambda}(\tau_f)^T &&= \frac{\partial \Phi}{\partial \mathbf{x}(\tau_f)} - \boldsymbol{\nu}^T \frac{\partial \boldsymbol{\phi}}{\partial \mathbf{x}(\tau_f)}
\end{aligned} \tag{6}$$

where $\boldsymbol{\nu} \in \mathbb{R}^q$ is Lagrange multiplier associated with the boundary condition $\boldsymbol{\phi}$. The optimality conditions of Eq. (6) define a *Hamiltonian boundary value problem* (HBVP).

4 Descriptions of the Pseudospectral Methods

The three pseudospectral methods in this comparison approximate the state using a finite basis of global polynomials, which are created by first discretizing the problem into N *nodes* along the trajectory. The locations of these nodes are unique to each method, along with the choice of basis polynomials that interpolates the trajectory between the nodes. Furthermore, all pseudospectral methods transcribe Eq. (2) into algebraic constraints, which are evaluated at a set of *collocation points*, $\mathcal{K}$. The set $\mathcal{K}$ is also unique to each method and may or may not be the same as the node set, $\mathcal{N}$, which is a subtle yet important distinction.

There are many different types of pseudospectral methods, but this comparison is limited to three methods: the Legendre, Radau, and Gauss pseudospectral methods, since these are the three methods with published costate approximation procedures. It is noted that the indexing in this section may be slightly different than the indexing contained in the references in order to create commonality between the three methods, and also to create NLPs that are equivalent in size. In all methods, N represents the total number of discretization points (nodes) used in the NLP. Moreover, in general, the index j is used to represent the j^{th} node, while the index k is used to represent the k^{th} collocation point.

4.1 Legendre Pseudospectral Method (LPM)

The Legendre pseudospectral method (LPM), like all pseudospectral methods, is based on approximating the state using a finite set of interpolating polynomials. The state is approximated using a basis of N Lagrange interpolating polynomials, $\mathcal{L}_i(\tau)$ $(i = 1, \ldots, N)$, as

$$\mathbf{x}(\tau) \approx \mathbf{X}(\tau) = \sum_{i=1}^{N} \mathcal{L}_i(\tau)\mathbf{X}(\tau_i) \tag{7}$$

In the LPM, the location of N nodes are described by the Legendre-Gauss-Lobatto (LGL) points on the interval $\tau \in [-1, 1]$. Note that the endpoints -1 and 1 are included in this state approximation.

Next, in the LPM, the LGL points are also the *collocation* points, meaning $\mathcal{N} = \mathcal{K}$. The derivative of Eq. (7) at the k^{th} LGL point, τ_k, is

$$\dot{\mathbf{x}}(\tau_k) \approx \dot{\mathbf{X}}(\tau_k) = \sum_{i=1}^{N} \dot{\mathcal{L}}_i(\tau_k)\mathbf{X}(\tau_i) = \sum_{i=1}^{N} D_{ki}\mathbf{X}(\tau_k), \quad (k = 1, \ldots, K) \tag{8}$$

where the differentiation matrix, $D \in \mathbb{R}^{K\times N}$, is defined as

$$D_{ki} = \begin{cases} \frac{P_{N-1}(\tau_k)}{P_{N-1}(\tau_i)}\frac{1}{\tau_k-\tau_i}, & \text{if } k \neq i \\ -\frac{(N-1)N}{4}, & \text{if } k = i = 1 \\ \frac{(N-1)N}{4}, & \text{if } k = i = N \\ 0, & \text{otherwise} \end{cases} \tag{9}$$

The continuous dynamics of Eq. (2) are then transcribed into the following set of K algebraic equations via collocation as follows:

$$\sum_{i=1}^{N} D_{ki}\mathbf{X}(\tau_i) - \frac{t_f - t_0}{2}\mathbf{f}(\mathbf{X}(\tau_k), \mathbf{U}(\tau_k)) = \mathbf{0}, \quad (k = 1, \ldots, K) \tag{10}$$

The approximated control, $\mathbf{U}(\tau)$, is defined in a similar manner as the state:

$$\mathbf{u}(\tau) \approx \mathbf{U}(\tau) = \sum_{i=1}^{N} \mathcal{L}_i(\tau)\mathbf{U}(\tau_i) \tag{11}$$

where $\tau_i,\ i = 1, \ldots, N$ are the LGL nodes. Note that this is not the only approximation to the control that can be used. Piecewise polynomial splines are also a common approximation to the control.

Next, the cost function of Eq. (1) is approximated simply as

$$J = \Phi(\mathbf{X}(\tau_N)) \tag{12}$$

where $\tau_N \equiv 1$ for the LPM. Lastly, the boundary constraint of Eq. (3) is expressed using τ_N and $\tau_1 \equiv -1$ as

$$\boldsymbol{\phi}(\mathbf{X}(\tau_1), \mathbf{X}(\tau_N)) = \mathbf{0} \tag{13}$$

The cost function of Eq. (12) and the algebraic constraints of Eqs. (10) and (13) define an NLP whose solution is an approximate solution to the continuous Mayer problem.

4.2 Radau Pseudospectral Method (RPM)

The Radau pseudospectral method has been primarily used in the chemical engineering community.[12,19] The location of the nodes in the Radau method are based on the flipped Legendre-Gauss-Radau (LGR) points, which lie on the interval $\tau \in (-1, 1]$. This set of points includes the final point but not the initial point. When discretizing optimal control problems, it is desirable that the discretization span the entire interval, including both endpoints. Therefore, in order to complete the full discretization of the time interval, the N discretization points are found using $N-1$ flipped Radau points plus the initial point, $\tau_0 \equiv -1$. The state approximation is constructed exactly the same as the Legendre pseudospectral method, using a basis of N Lagrange polynomials:

$$\mathbf{x}(\tau) \approx \mathbf{X}(\tau) = \sum_{i=0}^{N-1} \mathcal{L}_i(\tau)\mathbf{X}(\tau_i) \tag{14}$$

where Eq. (14) uses a slight index modification from $(i = 1, \ldots, N)$ in the LPM to $(i = 0, \ldots, N-1)$. This new index highlights the fact that the nodes $\tau_i,\ i = 0, \ldots, N-1$ are the initial point plus the $N-1$ LGR points.

Unlike the Legendre pseudospectral method, the Radau method uses a different number of collocation points, K, than discretization points, N. Specifically, the collocation points are the $N-1$ LGR points, while the discretization points are the LGR points plus the initial point, τ_0. Therefore, $K = N-1$ and $\mathcal{K} \subset \mathcal{N}$. The K collocation equations are then described as

$$\sum_{i=0}^{N-1} D_{ki}\mathbf{X}(\tau_i) - \frac{t_f - t_0}{2}\mathbf{f}(\mathbf{X}(\tau_k), \mathbf{U}(\tau_k)) = \mathbf{0}, \quad (k = 1, \ldots, K) \tag{15}$$

where τ_k are the LGR points and

$$\dot{\mathcal{L}}_i(\tau_k) = D_{ki} = \begin{cases} \dfrac{\dot{g}(\tau_k)}{(\tau_k - \tau_i)\dot{g}(\tau_i)}, & \text{if } k \neq i \\ \dfrac{\ddot{g}(\tau_i)}{2\dot{g}(\tau_i)}, & \text{if } k = i \end{cases} \tag{16}$$

where the function $g(\tau_i) = (1+\tau_i)[P_K(\tau_i) - P_{K-1}(\tau_i)]$ and τ_i, $(i = 0, \ldots, K)$ are the K LGR points plus the initial point. Since the collocation equations involve the control solely at the Radau points, the control is approximated using $N-1$ Lagrange polynomials, $\bar{\mathcal{L}}_k(\tau)$ $(k = 1, \ldots, N-1)$, as

$$\mathbf{u}(\tau) \approx \mathbf{U}(\tau) = \sum_{k=1}^{N-1} \bar{\mathcal{L}}_k(\tau)\mathbf{U}(\tau_k) \tag{17}$$

where τ_k again are the LGR points. This control approximation uses one less point than the state approximation, meaning there is no control defined at the initial point. See Table 1 for more details. In practice, the initial control is simply extrapolated from Eq. (17). The discretization of the rest of the optimal control problem is similar to the LPM. The continuous cost function of Eq. (1) is approximated as

$$J = \Phi(\mathbf{X}(\tau_K)) \tag{18}$$

Recall that in the Radau method, the final LGR point $\tau_K = 1$. Lastly, the boundary constraint of Eq. (3) is expressed as a function of the initial point and the final point as

$$\boldsymbol{\phi}(\mathbf{X}(\tau_0), \mathbf{X}(\tau_K)) = \mathbf{0} \tag{19}$$

The cost function of Eq. (18) and the algebraic constraints of Eqs. (15) and (19) define an NLP whose solution is an approximate solution to the continuous Mayer problem.

4.3 Gauss Pseudospectral Method (GPM)

The location of the nodes in the Gauss pseudospectral method are based on Legendre-Gauss (LG) points, which lie on the interval $\tau \in (-1, 1)$. This set of points includes neither the initial point nor the final point. Therefore, in order to fully discretize the time interval, the N discretization points are $(N-2)$ interior LG points, the initial point, $\tau_0 \equiv -1$, and the final point, $\tau_F \equiv 1$. The state is approximated using a basis of $N-1$ Lagrange interpolating polynomials, which is slightly smaller than the previous two methods,

$$\mathbf{x}(\tau) \approx \mathbf{X}(\tau) = \sum_{i=0}^{N-2} \bar{\mathcal{L}}_i(\tau)\mathbf{X}(\tau_i) \tag{20}$$

where τ_i $(i = 0, \ldots, N-2)$ are the initial point plus the $(N-2)$ LG points. Note that the final point, although it is part of the NLP discretization, is not part of the state approximation. This results in a state approximation that is one order less than the previous methods, which is necessary in order to have the equivalence property[24] between the KKT conditions and HBVP conditions. As in the Radau pseudospectral method, the Gauss pseudospectral method uses a different number of collocation points, K, than discretization points, N. Specifically, the collocation points are the LG points, while the discretization points are the LG points plus the initial point and final point. Therefore, $K = N-2$ and $\mathcal{K} \subset \mathcal{N}$. The K collocation equations are then described as

$$\sum_{i=0}^{N-2} D_{ki}\mathbf{X}(\tau_i) - \frac{t_f - t_0}{2}\mathbf{f}(\mathbf{X}(\tau_k), \mathbf{U}(\tau_k)) = \mathbf{0}, \quad (k = 1, \ldots, K) \tag{21}$$

where τ_k, $(k = 1, \ldots, K)$ are the LG points. Note that there is no collocation at either boundary. The control is approximated at the $(N-2)$ collocation points using a basis of $(N-2)$ Lagrange interpolating polynomials $\tilde{\mathcal{L}}_i(\tau)$, $(i = 1, \ldots, N-2)$ as

$$\mathbf{u}(\tau) \approx \mathbf{U}(\tau) = \sum_{i=1}^{N-2} \tilde{\mathcal{L}}_i(\tau)\mathbf{U}(\tau_i) \tag{22}$$

where τ_i, $(k = 1, \ldots, N-2)$ are the LG points. Table 1 lists all the Lagrange polynomial expressions described in this work.

An additional constraint must be added to the discretization in order to ensure that the the final state obeys the state dynamics. This is enforced by including a Gauss quadrature to approximate the integral of the dynamics across the entire interval:

$$\mathbf{X}(\tau_F) - \mathbf{X}(\tau_0) - \frac{t_f - t_0}{2}\sum_{k=1}^{K} w_k \mathbf{f}(\mathbf{X}(\tau_k), \mathbf{U}(\tau_k)) = \mathbf{0} \tag{23}$$

Table 1: Definitions of Lagrange interpolating polynomials.

# of Basis Polynomials	Application	Symbol	Indices
N	LPM ($\mathbf{x}$,$\mathbf{u}$,$\boldsymbol{\lambda}$)	$\mathcal{L}_i(\tau)$	$i = 1, \ldots, N$
N	RPM ($\mathbf{x}$)	$\mathcal{L}_i(\tau)$	$i = 0, \ldots, N-1$
$N-1$	RPM ($\mathbf{u}, \boldsymbol{\lambda}$)	$\mathcal{L}_i(\tau)$	$i = 1, \ldots, N-1$
$N-1$	GPM ($\mathbf{x}$,$\boldsymbol{\lambda}$)	$\mathcal{L}_i(\tau)$	$i = 0, \ldots, N-2$
$N-2$	GPM ($\mathbf{u}$)	$\tilde{\mathcal{L}}_i(\tau)$	$i = 1, \ldots, N-2$

where w_k are the Gauss weights and τ_k are the LG points. Lastly, the cost function is approximated by

$$J = \Phi(\mathbf{X}(\tau_F)) \tag{24}$$

and the boundary constraint is expressed as

$$\boldsymbol{\phi}(\mathbf{X}(\tau_0), \mathbf{X}(\tau_F)) = \mathbf{0} \tag{25}$$

The cost function of Eq. (24) and the algebraic constraints of Eqs. (21), (23), and (25) define an NLP whose solution is an approximate solution to the continuous Mayer problem.

5 Costate Estimation

All three aforementioned pseudospectral methods have established costate estimation procedures. In particular, a mapping between the KKT multipliers of the NLP and the costates of the continuous-time optimal control problem has been derived for all three methods. This section describes the costate estimation procedure for each method.

Legendre Pseudospectral Method

The KKT conditions of the NLP are derived by defining an augmented cost function, which brings the NLP constraints into the cost function via Lagrange multipliers. For LPM, the augmented cost is

$$J_a = \Phi(\mathbf{X}_N) - \tilde{\boldsymbol{\nu}}^T \boldsymbol{\phi}(\mathbf{X}_1, \mathbf{X}_N) - \sum_{k=1}^{K} \tilde{\boldsymbol{\Lambda}}_k^T \left(\sum_{i=1}^{N} D_{ki} \mathbf{X}_i - \frac{t_f - t_0}{2} \mathbf{f}(\mathbf{X}_k, \mathbf{U}_k) \right) \tag{26}$$

The KKT multipliers in Eq. (26) are $\tilde{\boldsymbol{\Lambda}}_k,\ (k = 1, \ldots, K)$, which are associated with the collocation equations of Eq. (10) and $\tilde{\boldsymbol{\nu}}$, which relates to the boundary condition of Eq. (13). A mapping from the KKT multipliers to the HBVP variables, given in Ref. 25, is stated as follows:

$$\begin{aligned} \boldsymbol{\Lambda}_k &= \frac{\tilde{\boldsymbol{\Lambda}}_k}{w_k}, \quad (k = 1, \ldots, K) \\ \boldsymbol{\nu} &= \tilde{\boldsymbol{\nu}} \end{aligned} \tag{27}$$

The HBVP variables $\boldsymbol{\Lambda}_k,\ (k = 1, \ldots, K)$, and $\boldsymbol{\nu}$, now defined in terms of NLP multipliers, can be substituted into the HBVP equations to determine if the optimality conditions have been met. It has been documented in the literature[26,28] that this mapping for the LPM does not provide an exact equivalence between the KKT conditions and the HBVP conditions. Specifically, the HBVP costate differential equations are not naturally collocated at the boundaries, meaning the boundary KKT conditions result in a linear combination of two HBVP conditions: the boundary costate dynamics and the transversality conditions. Although all the KKT conditions are not derived here, the conflicting KKT condition is shown, which is found by taking the partial derivative of the augmented cost function with respect to the initial state:

$$\sum_{k=1}^{K} \boldsymbol{\Lambda}_k^T D_{1k} + \frac{t_f - t_0}{2} \boldsymbol{\Lambda}_1^T \frac{\partial \mathbf{f}_1}{\partial \mathbf{X}_1} = \frac{1}{w_1} \left(-\boldsymbol{\Lambda}_1^T + \boldsymbol{\nu}^T \frac{\partial \boldsymbol{\phi}}{\partial \mathbf{X}_1} \right) \tag{28}$$

Notice that the left-hand side of the equation resembles the discretized costate dynamics and the right hand side resembles the transversality condition from the HBVP in Section 3. If these conditions were to map

exactly to the HBVP equations, each side of Eq. (28) should equal zero. However, Eq. (28) is inseparable in the LPM. Similarly, the KKT condition corresponding to the final state is

$$\sum_{k=1}^{K} \mathbf{\Lambda}_k^T D_{Nk} + \frac{t_f - t_0}{2} \mathbf{\Lambda}_N^T \frac{\partial \mathbf{f}_N}{\partial \mathbf{X}_N} = \frac{1}{w_N} \left(\mathbf{\Lambda}_N^T - \frac{\partial \Phi}{\partial \mathbf{X}_N} + \boldsymbol{\nu}^T \frac{\partial \boldsymbol{\phi}}{\partial \mathbf{X}_N} \right) \tag{29}$$

where again, the left-hand side is the discretized costate dynamics and the right-hand side is the transversality condition corresponding to the terminal state. It is impossible to get an exact mapping from the KKT conditions to the HBVP conditions without being able to separate these mixed conditions. Recent research[26] suggests that by applying relaxation techniques, linear independence constraint qualification, and an additional set of conditions called *closure conditions*, one can show that there exists a costate approximation that converges to the true costate in the limit as the number of collocation points increases to infinity. From Ref. 26, the most recent closure conditions are

$$\begin{aligned} \left\| -\mathbf{\Lambda}_1^* - \frac{\partial \Phi}{\partial \mathbf{x}_1}^T (\mathbf{X}_1^*, \mathbf{X}_N^*) - \left(\frac{\partial \boldsymbol{\phi}}{\partial \mathbf{x}_1} (\mathbf{X}_1^*, \mathbf{X}_N^*) \right)^T \boldsymbol{\nu}^* \right\|_\infty &\leq \delta_D \\ \left\| \mathbf{\Lambda}_N^* - \frac{\partial \mathbf{\Phi}}{\partial \mathbf{x}_N}^T (\mathbf{X}_1^*, \mathbf{X}_N^*) - \left(\frac{\partial \boldsymbol{\phi}}{\partial \mathbf{x}_N} (\mathbf{X}_1^*, \mathbf{X}_N^*) \right)^T \boldsymbol{\nu}^* \right\|_\infty &\leq \delta_D \end{aligned} \tag{30}$$

where $\mathbf{X}^*$, $\mathbf{\Lambda}^*$, and $\boldsymbol{\nu}^*$ represent a solution that satisfies both Eq. (30) and the set of KKT conditions. δ_D is a dual feasibility tolerance. These conditions are essentially the discretized HBVP transversality conditions.[1] The *covector mapping theorem*[26] states that there exists an N large enough so that the closure conditions and the KKT conditions are satisfied to a specified tolerance, δ. Upon close examination of Ref. 26 (and the references therein) any algorithm attempting to solve the combined KKT conditions and closure conditions must solve a mixed primal-dual feasibility problem. By including the closure conditions of Eq. (30) to form the "augmented optimality conditions", the discretized costate must now be included as a variables to be solved within the NLP. This procedure significantly increases the computational complexity in terms of both the NLP size and in the derivation of the equations themselves. In light of this, both LPM costate approximations (corresponding to Refs. 25 and 26) are considered in this work. The LPM costate is computed first according to Eq. (27), and then by formulating the primal-dual feasibility problem which attempts to satisfy the augmented KKT conditions, where the dual feasibility tolerance, δ_d, is set to the default tolerances of the NLP solver SNOPT. It is noted that improving the LPM costate approximation is an ongoing research topic, and better methods may be found in the future.

Similar to the LPM state approximation, the continuous-time costate approximation is represented by a basis of N Lagrange interpolating polynomials as

$$\boldsymbol{\lambda}(\tau) \approx \mathbf{\Lambda}(\tau) = \sum_{i=1}^{N} \mathbf{\Lambda}(\tau_i) \mathcal{L}_i(\tau) \tag{31}$$

Radau Pseudospectral Method

The augmented cost function corresponding to the Radau pseudospectral method is given as

$$J_a = \Phi(\mathbf{X}_K) - \tilde{\boldsymbol{\nu}}^T \boldsymbol{\phi}(\mathbf{X}_0, \mathbf{X}_K) - \sum_{k=1}^{K} \tilde{\mathbf{\Lambda}}_k^T \left(\sum_{i=0}^{K} D_{ki} \mathbf{X}_i - \frac{t_f - t_0}{2} \mathbf{f}(\mathbf{X}_k, \mathbf{U}_k) \right) \tag{32}$$

The KKT multipliers resulting from the NLP using the Radau pseudospectral method are $\tilde{\mathbf{\Lambda}}_k$, $(k = 1, \ldots, K)$, associated with the collocation equations of Eq. (15) and $\tilde{\boldsymbol{\nu}}$, associated with the boundary conditions of Eq. (19). The mapping from the KKT multipliers to the HBVP variables, given in Ref. 21, is stated as follows:

$$\begin{aligned} \mathbf{\Lambda}_k &= \frac{\tilde{\mathbf{\Lambda}}_k}{w_k}, \quad (k = 1, \ldots, K) \\ \boldsymbol{\nu} &= \tilde{\boldsymbol{\nu}} \end{aligned} \tag{33}$$

The K discrete costates $\mathbf{\Lambda}_k$, $(k = 1, \ldots, K)$, and the HBVP multiplier $\boldsymbol{\nu}$ are now defined in terms of the KKT multipliers. As in the LPM, it is understood that there is an imperfect match between the KKT

multipliers and the costate.[21] The Radau method has a similar conflict of HBVP equations when the partial derivative of the augmented cost function is taken with respect to the final state:

$$\sum_{k=1}^{K} \mathbf{\Lambda}_k^T D_{Kk} + \frac{t_f - t_0}{2} \mathbf{\Lambda}_K^T \frac{\partial \mathbf{f}_K}{\partial \mathbf{X}_K} = \frac{1}{w_K} \left(\mathbf{\Lambda}_K^T - \frac{\partial \Phi}{\partial \mathbf{X}_K} + \boldsymbol{\nu}^T \frac{\partial \boldsymbol{\phi}}{\partial \mathbf{X}_K} \right) \tag{34}$$

However, recent work[19,21] has demonstrated rapid convergence rates of the costate approximation using Eq. (33) as the number of nodes increases. The discrete values of the costate are used to construct a continuous-time approximation to the costate, which is represented by a basis of $N-1$ Lagrange interpolating polynomials, as in the control:[21]

$$\boldsymbol{\lambda}(\tau) \approx \mathbf{\Lambda}(\tau) = \sum_{i=1}^{K} \mathbf{\Lambda}(\tau_i) \bar{\mathcal{L}}_i(\tau) \tag{35}$$

where τ_k, $(k = 1, \ldots, K)$ are the LGR points. Note that since there is no collocation at the initial point, there is no initial costate estimate that is directly output from the NLP. However, the initial costate can be approximated by extrapolating Eq. (35).

Gauss Pseudospectral Method

The augmented cost function corresponding to the Gauss discretization of the Mayer problem is

$$\begin{aligned} J_a = \quad & \Phi(\mathbf{X}_F) - \tilde{\boldsymbol{\nu}}^T \boldsymbol{\phi}(\mathbf{X}_0, \mathbf{X}_F) - \sum_{k=1}^{K} \tilde{\mathbf{\Lambda}}_k^T \left(\sum_{i=0}^{K} D_{ki} \mathbf{X}_i - \frac{t_f - t_0}{2} \mathbf{f}(\mathbf{X}_k, \mathbf{U}_k) \right) \\ & - \tilde{\mathbf{\Lambda}}_F^T \left(\mathbf{X}_F - \mathbf{X}_0 - \frac{t_f - t_0}{2} \sum_{k=1}^{K} w_k \mathbf{f}(\mathbf{X}_k, \mathbf{U}_k) \right) \end{aligned} \tag{36}$$

The KKT multipliers resulting from the NLP using the Gauss pseudospectral method are $\tilde{\mathbf{\Lambda}}_k$ $(k = 1, \ldots, K)$, associated with the collocation equations of Eq. (21), $\tilde{\mathbf{\Lambda}}_F$, associated with the quadrature constraint of Eq. (23), and $\tilde{\boldsymbol{\nu}}$, associated with the boundary conditions of Eq. (25). The mapping from the KKT multipliers to the HBVP variables is:

$$\begin{aligned} \mathbf{\Lambda}_k &= \frac{\tilde{\mathbf{\Lambda}}_k}{w_k} + \tilde{\mathbf{\Lambda}}_F, \quad (k = 1, \ldots, K) \\ \boldsymbol{\nu} &= \tilde{\boldsymbol{\nu}} \,, \\ \mathbf{\Lambda}_f &= \tilde{\mathbf{\Lambda}}_F \end{aligned} \tag{37}$$

The variables $\mathbf{\Lambda}_k$, $\mathbf{\Lambda}_f$, and $\boldsymbol{\nu}$, now defined in terms of NLP multipliers, can be substituted into the HBVP equations to ensure the optimality conditions have been met. The discrete costate is then used to construct a continuous-time costate approximation, which is represented by a slightly different basis of $N-1$ Lagrange interpolating polynomials, defined as

$$\boldsymbol{\lambda}(\tau) \approx \mathbf{\Lambda}(\tau) = \sum_{i=1}^{K+1} \mathbf{\Lambda}(\tau_i) \bar{\mathcal{L}}_i(\tau) \tag{38}$$

where τ_i, $(i = 1, \ldots, K+1)$ are the LG points plus the final point, τ_F. Recall that the state is approximated using $(N-1)$ Lagrange polynomials based on the LG points and the *initial* time, while the costate, $\boldsymbol{\lambda}(\tau)$, is approximated using $(N-1)$ Lagrange polynomials consisting of the LG points and the *final* time. This discrepancy is required to preserve the mapping. As in the Radau method, an approximation for the initial costate, $\mathbf{\Lambda}_0$, cannot be pulled directly from a corresponding multiplier in the NLP because no such multiplier exists. However, an accurate approximation to the initial costate can be determined from the following equation:

$$\mathbf{\Lambda}_0 \equiv \tilde{\mathbf{\Lambda}}_F - \sum_{k=1}^{K} D_{k0} \tilde{\mathbf{\Lambda}}_k \tag{39}$$

which can be shown to satisfy the HBVP transversality condition

$$\mathbf{\Lambda}_0^T = \boldsymbol{\nu}^T \frac{\partial \boldsymbol{\phi}}{\partial \mathbf{X}_0} \tag{40}$$

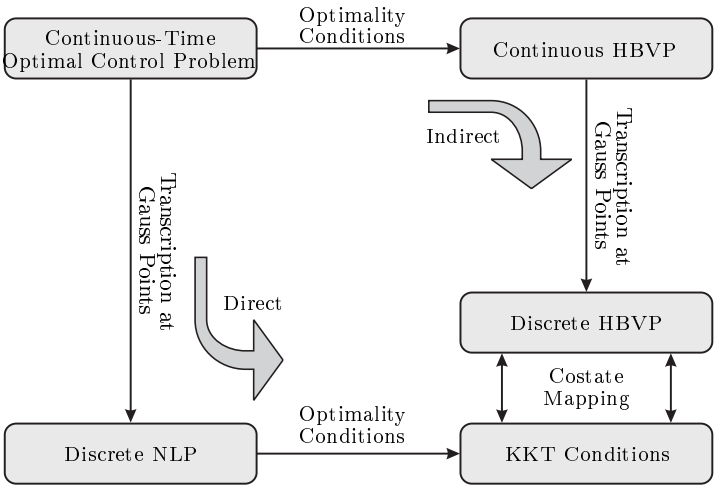

Figure 1: Direct vs Indirect Flow Diagram.

resulting in a perfect match between the HBVP equations and the KKT conditions. A diagram of the equivalence between the direct and indirect forms of the GPM is shown in Fig. 1 and specific details of this equivalence can be found in Ref.[24] As an alternative to Eq. (40), one could simply extrapolate Eq. (38) to τ_0, which would produce an equally accurate initial costate. All three methods for determining the initial costate are mathematically equivalent. The remainder of this work focuses on comparing these pseudospectral techniques in terms of their computational efficiency, approximation accuracy, and solution convergence rates with respect to the state, control, and especially the costate.

6 Comparison of Pseudospectral Methods

The three pseudospectral methods: GPM, RPM, and LPM, are compared on several examples of increasing complexity.

Example 1: Single State Problem

As a first example, consider the following one-dimensional optimal control problem. Minimize the cost functional

$$J = -y(t_f) \tag{41}$$

subject to the constraints

$$\dot{y}(t) = y(t)u(t) - y(t) - u^2(t) \tag{42}$$

and the initial condition

$$y(t_0) = 1 \tag{43}$$

where $y(t)$, is the state, $u(t)$ is the control, and $t_f = 5$ is the final time. The exact solution is

$$y^*(t) = \frac{4}{1+3e^t} \tag{44}$$

$$\lambda^*(t) = ae^{(2\ln(1+3e^t)-t)} \tag{45}$$

$$u^*(t) = .5y^*(t) \tag{46}$$

$$\text{where } a = \frac{-1}{(e^{-5}+6+9e^5)} \tag{47}$$

$$\tag{48}$$

$\lambda^*(t)$ is the costate associated with the optimal solution. The exact solution is depicted in Fig. 2.

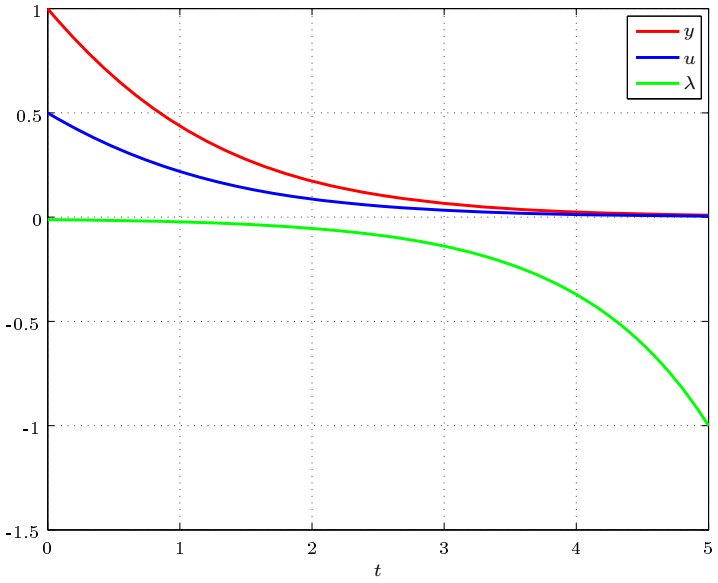

Figure 2: Exact solution for the single state example

Solution Using 10 Nodes

This problem is solved by all three pseudospectral techniques mentioned previously. Each method transcribes the optimal control problem into an NLP that consists of 10 nodes ($N = 10$). The resulting NLP is then solved with SNOPT[40] using the default feasibility and optimality tolerances (10^{-6}). In Fig. 3, the absolute value of the error between the NLP state and the exact solution ($|Y(\tau_i) - y^*(\tau_i)|$, $i = 1, \ldots, N$) is plotted for each node and all three approaches. Note that the error in initial state is not included since SNOPT satisfies state equality bounds (i.e. the initial conditions) to double precision (10^{-16}). When comparing the state accuracy of the three methods in this example, there is no clear method that outperforms the other two across the entire interval. In fact, the only notable difference in Fig. 3 is that both the GPM and RPM formulations produce an extremely accurate terminal state. It is hypothesized that the highly accurate terminal state for the GPM is due to the additional quadrature constraint in the NLP.

The control error is plotted in Fig. 4. As described in Section 4, each method approximates the control using a different number of points. In the LPM the control is discretized at all N discretization points, while Radau discretizes the control at the $(N-1)$ Radau points, and the Gauss method discretizes the control at the $(N-2)$ Gauss points. In the Radau discretization, the initial control is missing. Similarly, in the Gauss discretization, the control at both the initial and final time are missing. These missing control values can be computed by a variety of techniques. The approach often outlined in the literature[6,19] is to extrapolate the control approximation equations, Eqs. (11), (17) and (22). This technique was used to produce the boundary controls in Fig. 4, but it is noted that spline extrapolations[41] are often used in practice as well. Intuition suggests that extrapolated controls would be less accurate than the control values drawn from the NLP at the collocation points. In Fig. 4, the estimate for the boundary control using the GPM is indeed less accurate than the control at the collocation points. Likewise, the accuracy of the extrapolated initial control for the RPM is equally poor. Interestingly, the extrapolated controls for the GPM and RPM are still less than the largest control error in the LPM.

Perhaps a more appropriate way to measure control error is to propagate the state dynamics according to the NLP control and examine the change in cost. In this comparison, the state dynamics are propagated using the MATLAB function ODE45, where the control approximations are represented using Eqs. (11), (17), and (22). The resulting propagated state error is shown in Fig. 5. Recall that the costs is equal to the final state, $-y(t_f)$, meaning the LPM clearly has the largest error in optimal cost based on its propagated control.

The largest discrepancy between the methods is seen in the costate comparison of Fig. 6. The LPM costate approximation is several orders of magnitude worse than the other two methods across all values of N. This example highlights a common problem[26] with the original LPM costate approximation[25] where the approximation tends to "wiggle" about the true solution, shown clearly in Fig. 7. As suggested in Section 5,

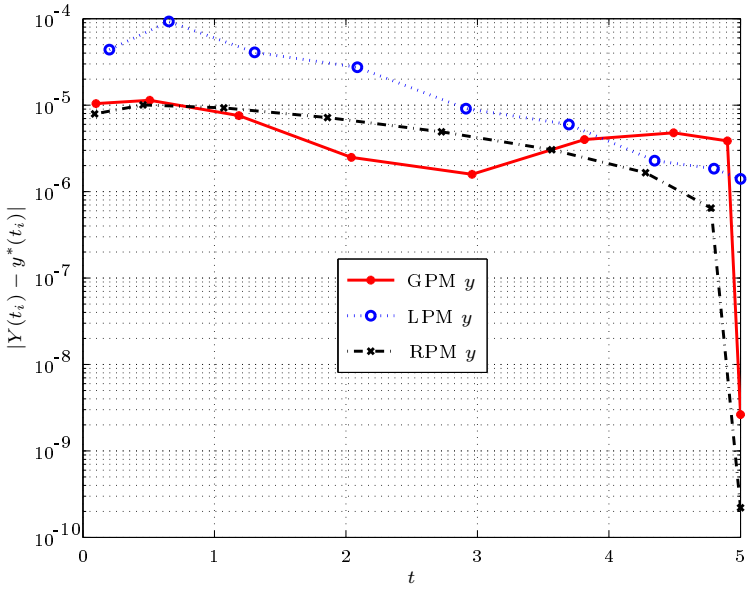

Figure 3: State error for the single state example.

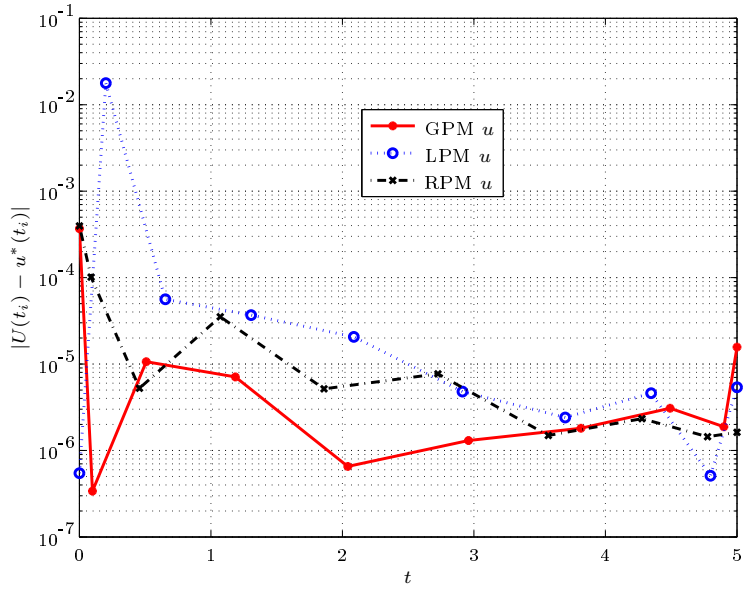

Figure 4: Control error for the single state example.

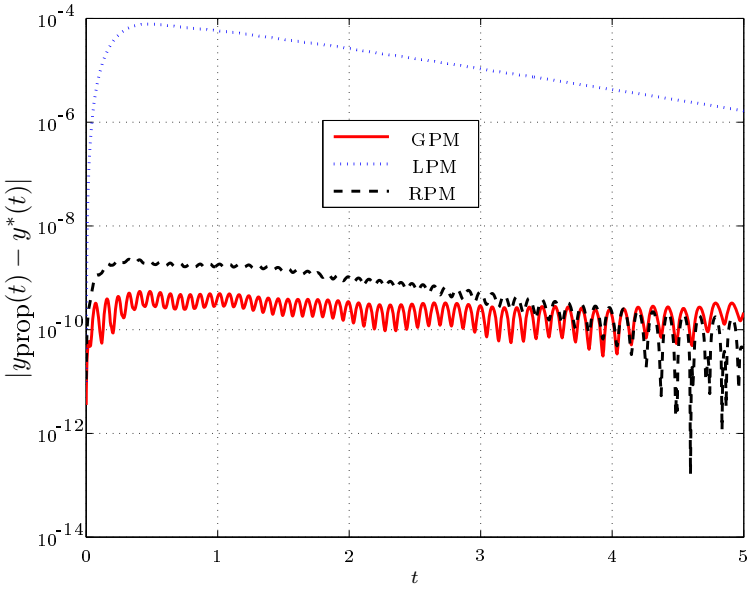

Figure 5: Error in the propagated state for the single state example.

the largest error in the LPM costate occurs at the boundaries due to the conflicting constraints. The GPM, which has no constraint conflict, provides extremely accurate boundary costates, while the RPM produces an accurate costate everywhere but the final costate (likely due to the conflict explained in Section 5 at the terminal time).

Convergence

In pseudospectral methods, like most direct methods, the solution accuracy can be improved by increasing the number of nodes used to discretize the problem.[11] The rate of convergence to the true optimal solution is extremely important as it can help determine the number of nodes needed to adequately approximate the solution. The convergence rates of the state and costate are shown in Figs. 8 and 9, respectively. In these figures, each method was solved repeatedly while the number of nodes was increased from 5 to 50. The error shown in these figures is the maximum absolute error over all the nodes ($||X(\tau_i) - x^*(\tau_i)||_\infty$, $\forall \tau_i \in [-1, 1]$). As seen in Figure 8, the state convergence rate for all three methods is quite similar. The steep convergence rate depicts the "spectral" convergence that is characteristic of pseudospectral methods.[8] Naturally, once the error drops below the tolerances of the NLP, the error stops improving. The convergence rate for the costate is shown in Fig. 9. The Gauss and Radau methods show rapid convergence rates for the costate, which even outperform the state. Fig. 9 also shows the apparent lack of convergence for the costate using the LPM. It is clear that increasing the number of nodes does not improve the costate error by any significant amount for this example. As mentioned previously, Gong and Ross have devised a modified covector mapping theorem[26] to improve the costate approximation for the LPM. This covector mapping theorem includes a set of "closure conditions" that, when used in conjunction with the original NLP constraints, creates a mixed primal-dual feasibility problem where the variables are the state, control, costate and Lagrange multipliers associated with boundary constraints. When this augmented NLP is posed and solved, the resulting costate convergence rate in Fig. 10 is indeed much better, although not as rapid as either the Gauss or Radau methods. As mentioned previously, this increased accuracy comes at a significant computational burden. The number of NLP variables almost doubles, and intimate knowledge of optimal control theory is necessary to properly formulate the closure conditions. Due to the large number of extra steps necessary to produce this result, further computation of LPM costates in this paper involves only the original costate estimation procedure of Eq. (27), since this is the likely method to be used in practice and more closely resembles the techniques of the other two methods.

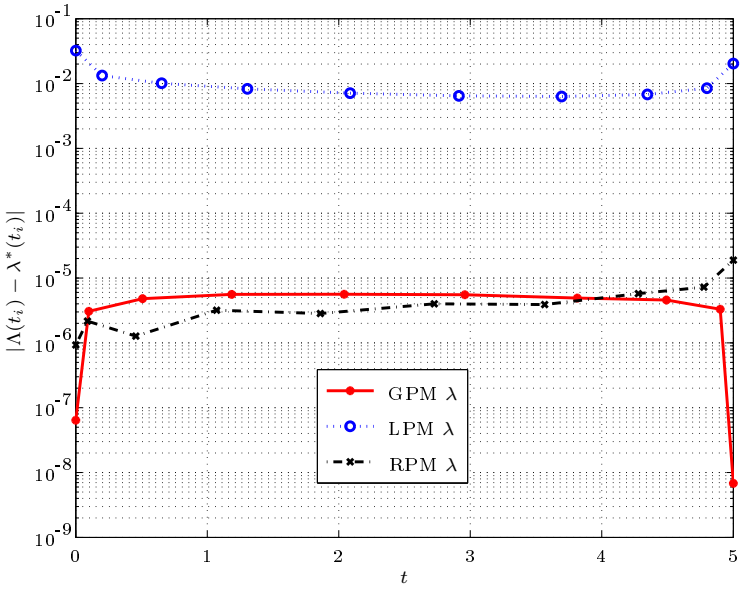

Figure 6: Costate error for the single state example.

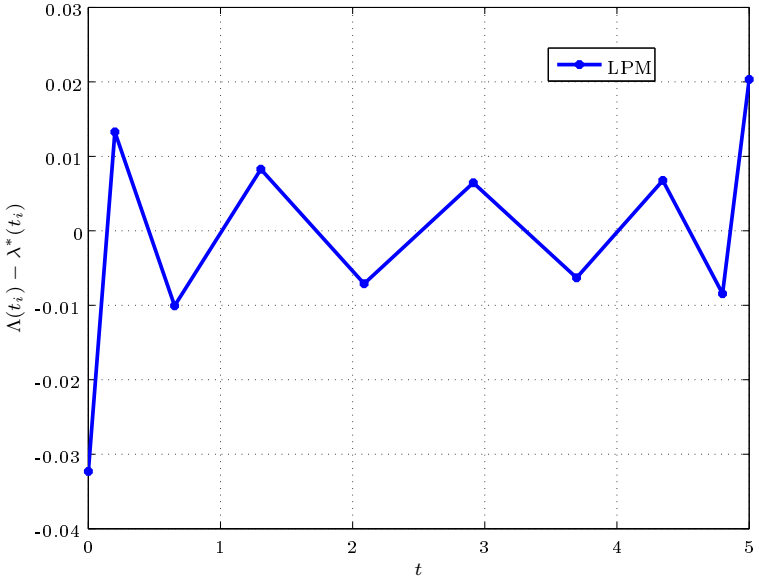

Figure 7: LPM costate error for the single state example.

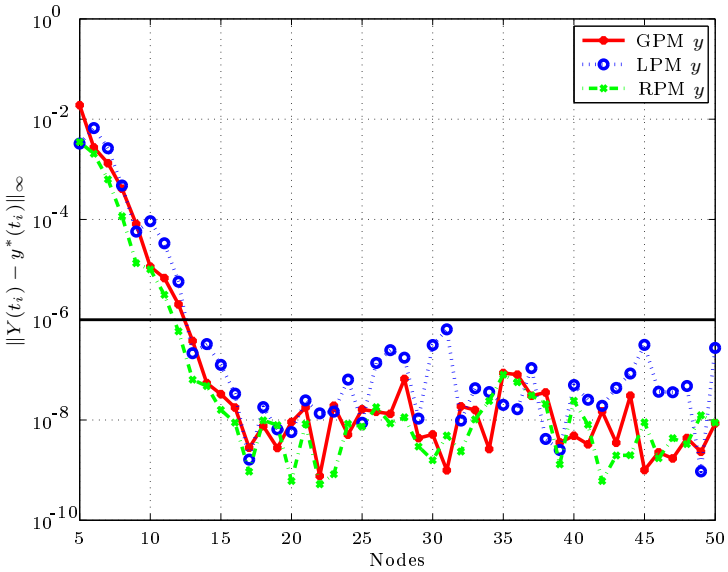

Figure 8: State convergence for the single state example

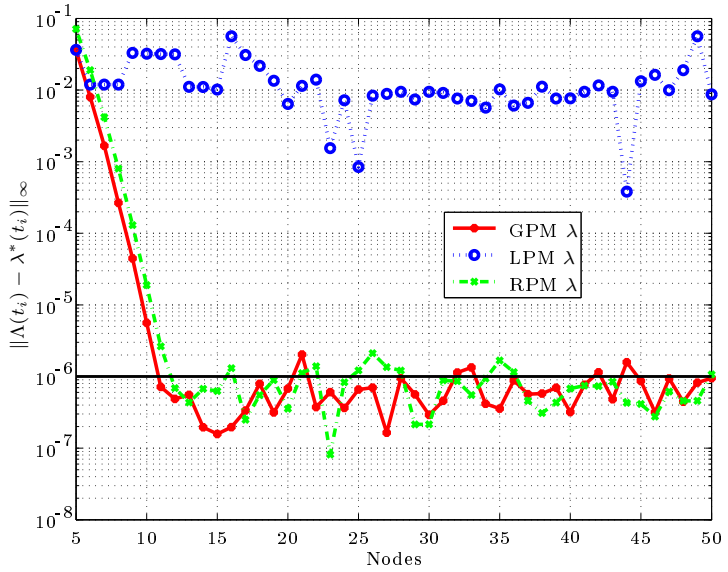

Figure 9: Costate convergence for the single state example

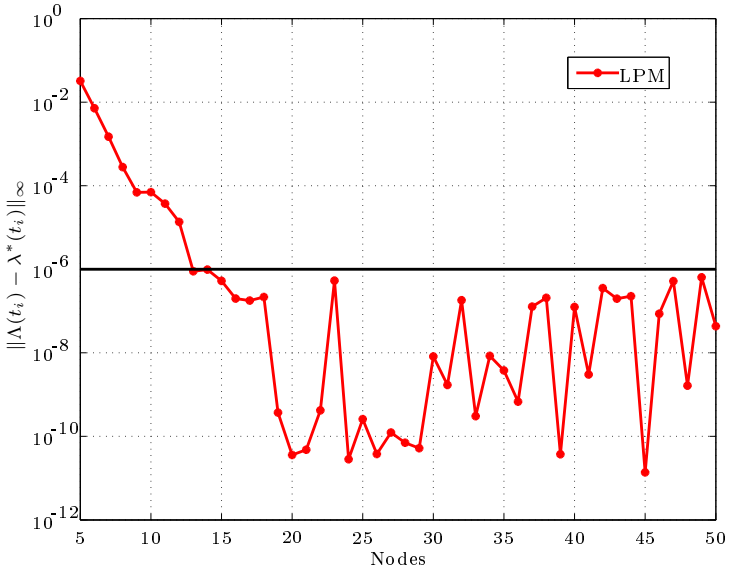

Figure 10: Improved LPM costate convergence using closure conditions and solving mixed primal-dual feasibility problem for the single state example.

Example 2: Two State Problem

As a second example, consider the following two-dimensional optimal control problem. Minimize the cost functional

$$J = y_2(t_f) \tag{49}$$

subject to the dynamic constraints

$$\dot{y}_1(t) = 0.5y_1(t) + u(t) \tag{50}$$
$$\dot{y}_2(t) = y_1^2(t) + 0.5u^2(t) \tag{51}$$

and the boundary conditions,

$$y_1(t_0) = 1 \quad y_1(t_f) = 0.5 \tag{52}$$
$$y_2(t_0) = 0 \tag{53}$$

where $t_f = 5$. Note that this problem contains a terminal bound on the first state, $y_1(t_f)$. The exact solution is of the form

$$y_1^*(t) = a_1 e^{1.5t} + a_2 e^{-1.5t}, \tag{54}$$
$$y_2^*(t) = a_3(e^{1.5t})^2 - a_4(e^{-1.5t})^2 + c_1, \tag{55}$$
$$\lambda_1^*(t) = a_5 e^{1.5t} + a_6 e^{-1.5t}, \tag{56}$$
$$\lambda_2^*(t) = 1, \tag{57}$$
$$u^*(t) = -\lambda_1^*(t) \tag{58}$$

The exact solution can be seen in Fig. 11.

10 Node Solution

Again, all three methods are compared using 10 nodes. The error plots are formulated in the same manner as the previous example problem. Fig. 12 shows the state error at each node for all three methods. Recall that the error for the states involving the boundary conditions is $O(10^{-16})$, and therefore outside the range of the plot. As in the single-state example, there is no method that clearly outperforms the other two in state accuracy, although Fig. 12 may suggest that the RPM is slightly more accurate that the other two methods.

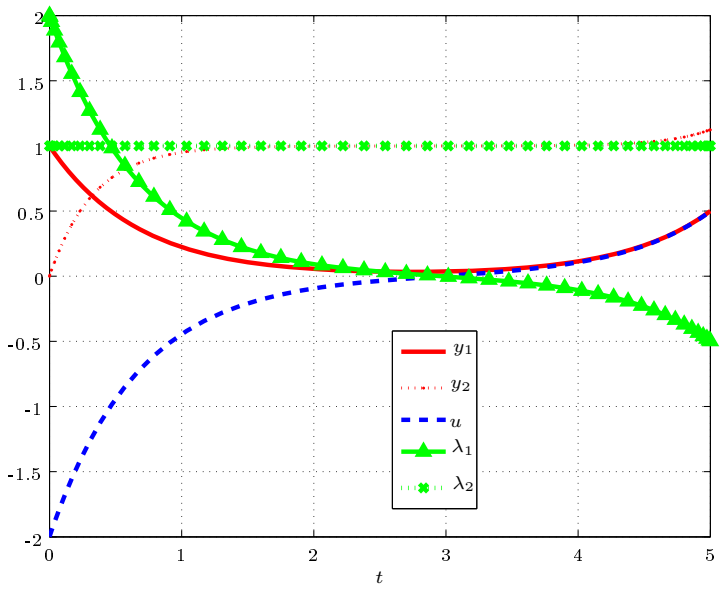

Figure 11: Exact solution for the two state example.

Moreover, the GPM and Radau formulations again show an increase in accuracy for the unconstrained terminal state, $y_2(t_f)$.

The control error is plotted in Fig. 13. For this problem, the LPM control estimate is worse than both GPM and RPM at all points. Like the previous example, the NLP control is used to propagate the state equations, and the resulting state accuracy is compared in Fig. 14 for both states. In this example, the RPM produces the most accurate propagated solution, followed by the GPM, and lastly LPM.

Reconfirming the results in the first example, the largest discrepancy between the methods is seen in the costate comparison, and specifically with the second costate, shown in Fig. 15. The LPM method produces the worst costate approximation for both λ_1 and λ_2. Although it is not shown, λ_2 from the LPM exhibits the same "wiggle" phenomenon around the true costate as seen in the first example. In terms of boundary costates, the GPM method provides the most accurate boundary costates, and the Radau method provides a relatively accurate initial costate, but an inaccurate final costate, as expected from the theoretical analysis.

Convergence

The convergence rates of the state and costate are shown in Figs. 16 and 17, respectively. Each approach was solved repeatedly while the number of nodes in the problem was increased from 5 to 30. The error shown in these figures is computed in the same manner as the previous example. Fig. 16 displays both states, y_1 and y_2, and there is a noticeable difference between the convergence rate of y_1 between the LPM and the other two methods. The state convergence rate of the RPM is the fastest for this example. It is hypothesized that this slight deficiency in the GPM state approximation is due to the fact that the GPM state approximation uses one less point in its formulation. Similar convergence results are seen in the costate convergence plot of Fig. 17, but now the GPM is slightly better than the RPM for λ_1. The second costate is constant in this example, and both the GPM and RPM approaches can approximate this first order polynomial quite well, even using only a few nodes, as expected with pseudospectral methods. However, the LPM approach displays the wiggle phenomenon and thus has a much slower convergence rate for λ_2. In terms of the first costate, the GPM has the fastest convergence rate.

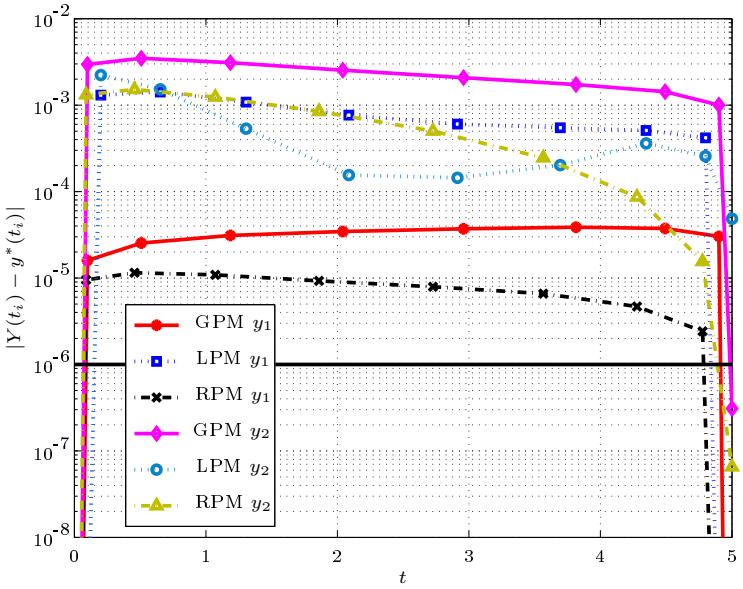

Figure 12: State error for the two state example

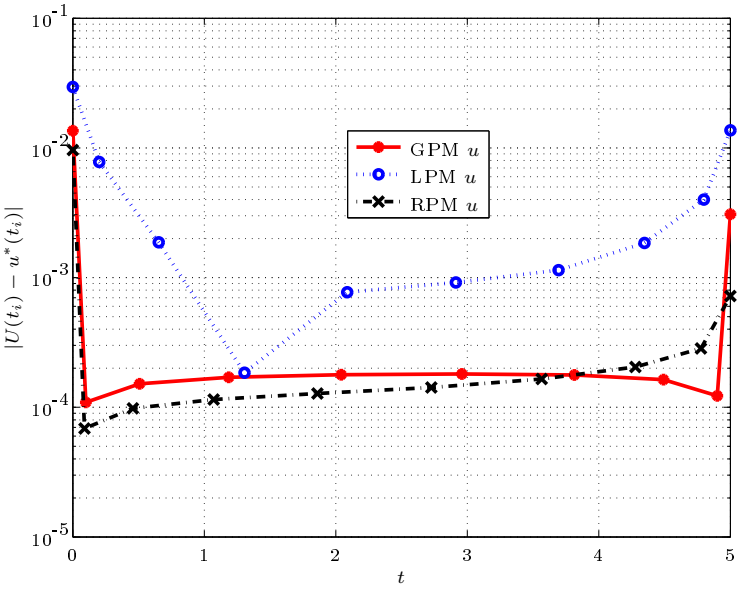

Figure 13: Control error for the two state example.

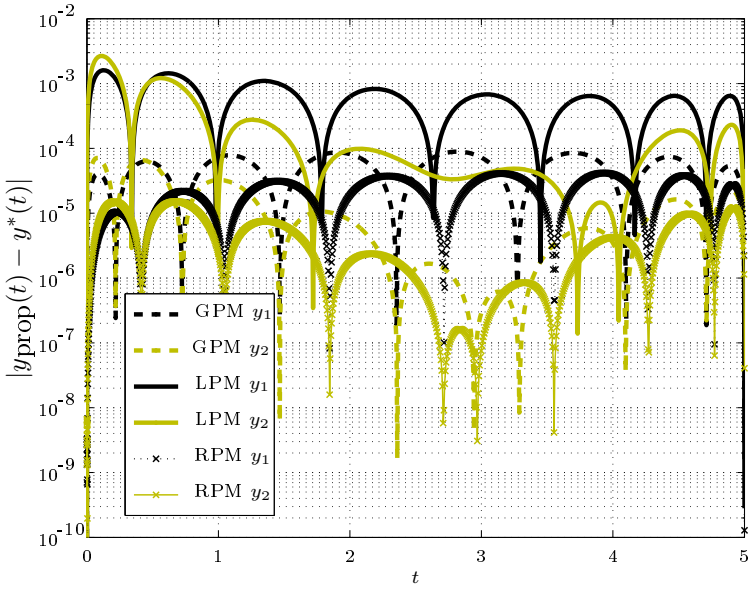

Figure 14: Error in the propagated state for the two state example

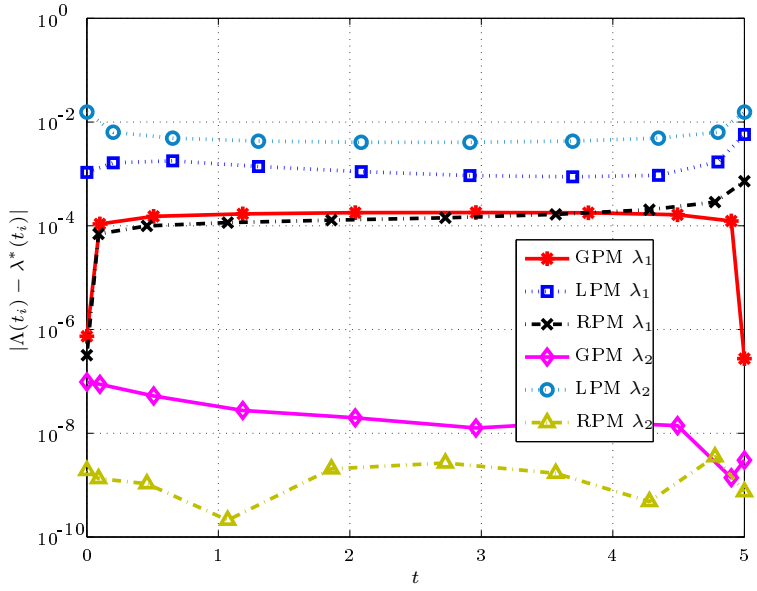

Figure 15: Costate error for the two state example.

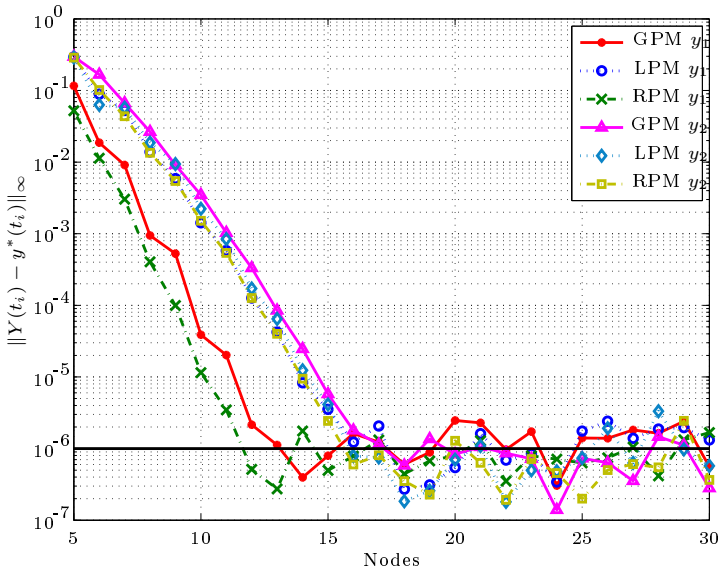

Figure 16: State convergence for the two state example.

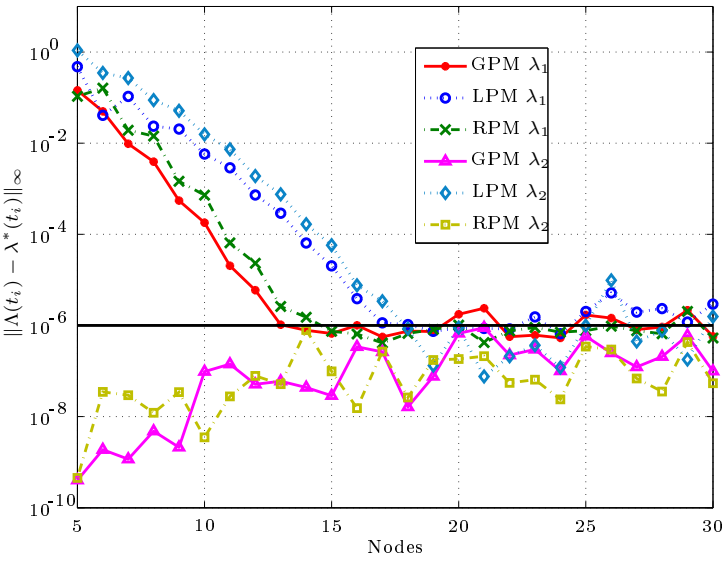

Figure 17: Costate convergence for the two state example

6.1 Example 3: Orbit-Raising Problem

Consider the problem of transferring a spacecraft from an initial circular orbit to the largest possible circular orbit in a fixed time using a low-thrust engine. The state for this problem is composed of the spacecraft radial position, $r(t)$, radial velocity, $u(t)$, and tangential velocity, $v(t)$. The control, $\phi(t)$, is the angle between the thrust vector and the tangential velocity. The optimal control problem is stated as follows. Minimize the cost functional

$$J = -r(t_f) \tag{59}$$

subject to the dynamic constraints

$$\begin{aligned} \dot{r} &= u \\ \dot{u} &= \frac{v^2}{r} - \frac{\mu}{r^2} + \frac{T\sin\phi}{m_0 - |\dot{m}|t} \\ \dot{v} &= -\frac{uv}{r} + \frac{T\cos\phi}{m_0 - |\dot{m}|t} \end{aligned} \tag{60}$$

and the boundary conditions

$$\begin{array}{llll} r(0) = 1, & & r(t_f) = \text{free} \\ u(0) = 0, & & u(t_f) = 0, \\ v(0) = \sqrt{\dfrac{\mu}{r(0)}}, & & v(t_f) = \sqrt{\dfrac{\mu}{r(t_f)}}, \end{array} \tag{61}$$

where μ is the gravitational parameter, T is the thrust magnitude, m_0 is the initial mass, and $\dot{m}$ is the fuel consumption rate. These parameters are given as follows in normalized units:

$$T = 0.1405, \quad \dot{m} = 0.0749, \quad m_0 = \mu = 1, \quad t_f = 3.32 \tag{62}$$

This problem does not have an analytic solution, but it has been solved numerically many times[38,43] so its solution is well known. The results of each method was compared against the solution to a boundary value problem solution using the MATLAB function BVP4C with a tolerance of 10^{-9}. For brevity, an error analysis is presented for only the first state and costate, $r(t)$ and $\lambda_r(t)$. The remaining state and costate errors look very similar to the one presented. Fig. 18 shows the state error throughout the trajectory. The RPM has a slightly better state accuracy than the other two methods, although they all have very similar maximum state errors. Interestingly, in Fig. 19, the RPM also has a slightly better costate approximation than the other two methods and does not exhibit the usual reduced costate accuracy at the terminal time. This is perhaps due to the fact that the optimal control for this problem has a discontinuity, which has adverse affects on pseudospectral methods. As in the previous examples, the LPM costate convergence rate is significantly worse than the other two methods, shown in Fig. 20. Note that the convergence rates for all three methods are slower than in the previous examples due to the discontinuity in the optimal control profile, however the relative performance is still captured and is consistent with the previous examples.

7 Rationale for Choosing a Method

Based on the results of all three example problems and a detailed analysis of the mathematics of each method, there may be certain circumstances under which each method should be chosen. Inaccurate boundary costates in the LPM method are attributed to boundary costs or constraints in the problem. However, if the original problem has no boundary constraints or costs, then the conflicting HBVP equations disappear and LPM would be an appropriate method. The RPM costate inaccuracy only occurs at the terminal time, so problems without terminal constraints would likely mitigate the errors in the final costate. The GPM has a perfect mapping with the HBVP equations, and therefore creates accurate costate estimates for general problems with both initial and final constraints.

Lastly, Table 2 presents NLP computation times for all three examples and all three methods. The results listed correspond to a Pentium 4, 3.2 GHz machine, where the constraint Jacobian was computed using numerical derivatives. All three methods have very similar computation times, which is expected since the methods are very similar in terms of their problem density. Note that the codes used in this analysis are not optimized for computational efficiency. Streamlined coding, analytic derivatives (or automatic differentiation), and a transition to faster programming languages like C or Fortran would significantly improve the computation times of all three methods.

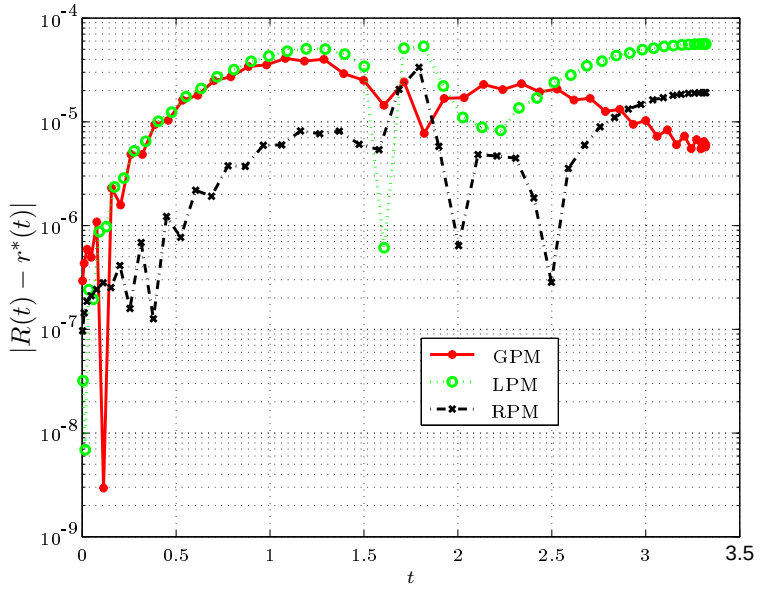

Figure 18: Error in state $r(t)$ for the orbit raising problem

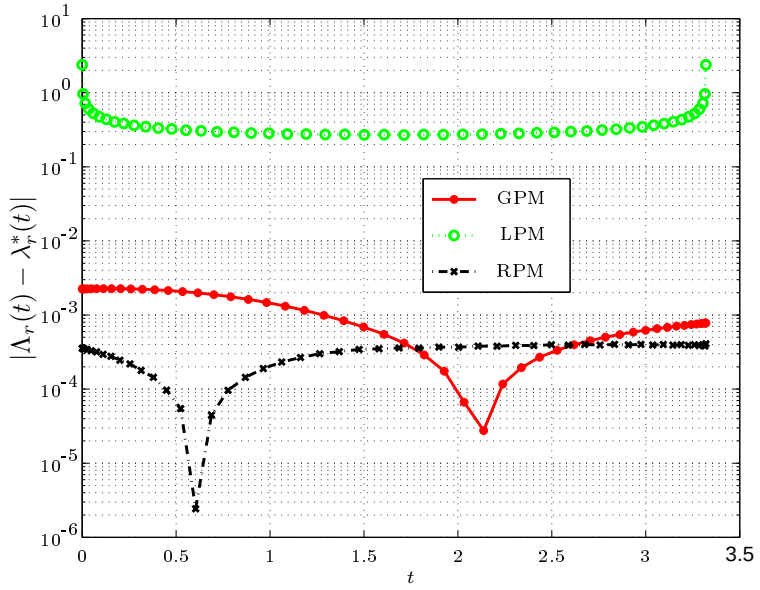

Figure 19: Error in the costate $\lambda_r(t)$ for the orbit raising problem

Table 2: CPU Times for each example and each method, in seconds.

	GPM	RPM	LPM
Example 1 ($N = 50$)	4.265	4.203	4.750
Example 2 ($N = 30$)	2.203	2.296	2.688
Example 3 ($N = 40$)	3.969	4.046	5.156

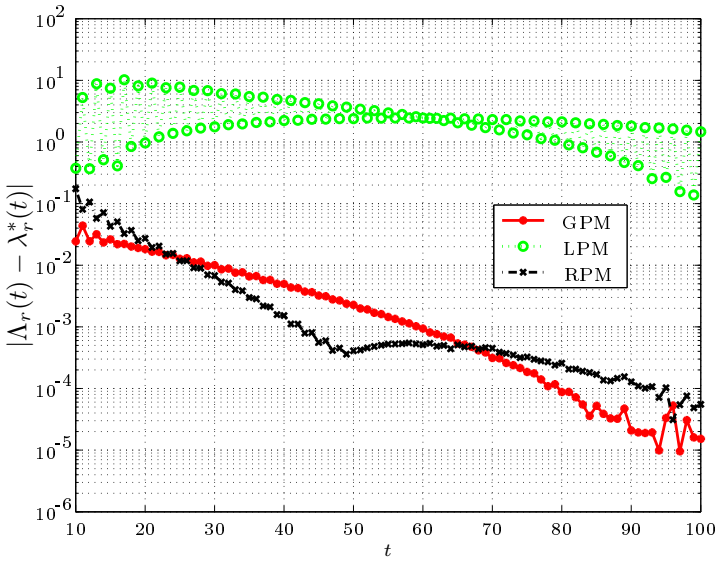

Figure 20: Convergence of costate λ_r for the orbit raising problem

8 Conclusions

A comparison has been made between three established pseudospectral methods that have been used in recent years in the numerical solution of optimal control problems. In particular, the Legendre, Radau, and Gauss pseudospectral methods have been compared in terms of the accuracy of the state, control, and costate. Three examples are used in the study to identify the key differences between the three methods. The results of this study indicate that the accuracy of the Radau and Gauss methods are very similar in accuracy, while both of these methods significantly outperform the traditional Legendre method in terms of costate accuracy. Furthermore, it is found that the computational efficiency of the three methods is quite comparable. Based on these results and a detailed analysis of the mathematics of each method, a rationale has been constructed to determine when each method should be chosen.

Acknowledgments

The authors gratefully acknowledge support for this research by The Charles Stark Draper Laboratory, Inc.

References

[1]Kirk, D. E. , *Optimal Control Theory*, Prentice-Hall, Englewood Cliffs, NJ, 1970.

[2]Pontryagin, L.S., Boltyanskii, V., Gamkrelidze, R., Mischenko, E., *The Mathematical Theory of Optimal Processes*. New York: Interscience, 1962.

[3]Betts, John T., "Survey of Numerical Methods for Trajectory Optimization," *Journal of Guidance, Control, and Dynamics*, Vol. 21, No. 2, March-April 1998

[4]Hull, D. , "Conversion of Optimal Control Problems into Parameter Optimization Problems," AIAA Guidance, Navigation, and Control Conference, San Diego, CA, July 29-31, 1996

[5]Vasantharajan, S. , and Biegler, L. T. , "Simultaneous Strategies for Optimization of Differential-Algebraic Systems with Enforcement of Error Criteria," *Computers and Chemical Engineering*, Vol. 14, No. 10, pp.1083-1100, 1990.

[6]Elnagar, G., Kazemi, M., Razzaghi, M., "The Pseudospectral Legendre Method for Discretizing Optimal Control Problems," *IEEE Transactions on Automatic Control*, Vol. 40, No. 10, October 1995

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

[8]Fornberg, B., *A Practical Guide to Pseudospectral Methods*, Cambridge University Press, 1998.

[9]Villadsen, J. , and Michelsen, M. , "Solution of differential equation models by polynomial approximation," Prentice Hall, New Jersey, 1978.

[10]Cuthrell, J. E. , and Biegler, L. T. , "On the Optimization of DIfferential-Algebraic Process Systems," *AIChE Journal*, Vol.33, No.8, pp.1257-1270, Aug. 1987.

[11]Trefethen, L. N., *Spectral Methods in MATLAB*, SIAM Press, Philadelphia, 2000.

[12]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, pp.49–62, 1989.

[13]Elnagar, G. E. , and Kazemi, M. A. , "Pseudospectral Legendre-based optimal computation of nonlinear constrained variational problems," *Journal of Computational and Applied Mathematics*, Vol. 88, pp.363-375, 1997.

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

[15]Ross, I. M. , and Fahroo, F. , "A Direct Method for Solving Nonsmooth Optimal Control Problems", *Proceedings of the 2002 World Congress of the International Federation on Automatic Control*, IFAC, Barcelona, July 2002.

[16]Rao, A. V. , "Extension of a Pseudospectral Legendre Method to Non-Sequential Multiple-Phase Optimal Control Problems", *AIAA Guidance, Navigation, and Control Conference*, AIAA Paper No. 2003-5634, Austin, TX, Aug. 11-14, 2003.

[17]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.

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

[19]Kameswaran, S. and Biegler, L. T., "Convergence Rates for Dynamic Optimization Using Radau Collocation," *SIAM Conference on Optimization*, Stockholm, Sweden, 2005.

[20]Fahroo, F. and Ross, I., "Pseudospectral Methods for Infinite Horizon Nonlinear Optimal Control Problems," *2005 AIAA Guidance, Navigation, and Control Conference*, AIAA Paper 2005-6076, San Francisco, CA, August 15–18, 2005.

[21]Kameswaran, S. and Biegler, L. T., "Convergence Rates for Direct Transcription of Optimal Control Problems at Radau Points," *Proceedings of the 2006 American Control Conference*, Minneapolis, Minnesota, June 2006.

[22]Vlassenbroeck, J. and Van Doreen, R., "A Chebyshev Technique for Solving Nonlinear Optimal Control Problems," *IEEE Transactions on Automatic Control*, Vol. 33, No. 4, 1988, pp. 333–340.

[23]Vlassenbroeck, J., "A Chebyshev Polynomial Method for Optimal Control with State Constraints," *Automatica*, Vol. 24, 1988, pp. 499–506.

[24]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, 2006, pp. 1435–1440.

[25]Fahroo, F. and Ross, I. M., "Costate Estimation by a Legendre Pseudospectral Method," *Journal of Guidance, Control, and Dynamics*, Vol. 24, No. 2, March-April 2002, pp. 270-277.

[26]Gong, Q. , Ross, I. , Kang, W. , and Fahroo, F. , "On the Pseudospectral Covector Mapping Theorem for Nonlinear Optimal Control," *IEEE Conference on Decision and Control*, Dec 13-15, 2006.

[27]Williams, P. , "Jacobi Pseudospectral Method for Solving Optimal Control Problems", *Journal of Guidance*, Vol. 27, No. 2,2003

[28]Williams, P. , "Hermite-Legendre-Gauss-Lobatto Direct Transcription Methods In Trajectory Optimization," Advances in the Astronautical Sciences. Vol. 120, Part I, pp. 465-484. 2005

[29]Williams, P. , "A Comparison of Differentiation and Integration Based Direct Transcription Methods," AAS/AIAA Space Flight Mechanics Conference, Copper Mountain, CO, Jan. 23-27, 2005. Paper AAS 05-128.

[30]Fahroo, F. and Ross, I. M., "Direct Trajectory Optimization by a Chebyshev Pseudospectral Method," *Journal of Guidance, Control, and Dynamics*, Vol. 25, No. 1, January-February 2002, pp. 160–166.

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

[32]Paris, S. W. , Riehl, J. P. , Sjuaw, W. .K. , "Enhanced Procedures for Direct Trajectory Optimization Using Nonlinear Programming and Implicit Integration," *AIAA/AAS Astrodynamics Specialist Conference and Exhibit*, Keystone CO, Aug 2006

[33]Fahroo, F. , Ross, I. M. , "On Discrete-Time Optimality Conditions for Pseudospectral Methods," *AIAA/AAS Astrodynamics Specialist Conference and Exhibit*, Keystone, CO, Aug 2006.

[34]Hull, D. G., *Optimal Control Theory for Applications*, Springer-Verlag, New York, 2003.

[35]Davis, P., *Interpolation & Approximation*, Dover Publications, 1975.

[36]Davis, P., Rabinowitz, P., *Methods of Numerical Integration*, Academic Press, 1984.

[37]Williams, P. , "A Gauss-Lobatto Quadrature Method For Solving Optimal Control Problems," Australia Mathematical Society, ANZIAM J.47, pp.C101-C115, July, 2006.

[38]Benson, D., *A Gauss Pseudospectral Transcription for Optimal Control*, Ph.D. Dissertation, Department of Aeronautics and Astronautics, Massachusetts Institute of Technology, November 2004.

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

[40]Gill, P.E., Murray, W., and Saunders, M.A., "SNOPT: An SQP Algorithm for Large Scale Constrained Optimization," *SIAM Journal on Optimization*, Vol. 12, No. 4, 2002

[41]Betts, John T., *Practical Methods for Optimal Control Using Nonlinear Programming* , SIAM. (2001)

[42]Huntington, G. T., Benson, D. A., and Rao, A. V., "Post-Optimality Evaluation and Analysis of a Formation Flying Problem via a Gauss Pseudospectral Method," *Proceedings of the 2005 AAS/AIAA Astrodynamics Specialist Conference*, AAS Paper 05-339, Lake Tahoe, California, August 7–11, 2005.

[43]Bryson, A., Ho, Y., *Applied Optimal Control*, Blaisdell Publishing Co. (1969)