

# Feasibility of the Galerkin Optimal Control Method

Randy Boucher, Wei Kang and Qi Gong

**Abstract**—A new numerical technique is presented for solving optimal control problems. This paper introduces a direct method that calculates optimal trajectories by discretizing the system dynamics using Galerkin numerical techniques and approximates the cost function with quadrature. We show that the Galerkin optimal control method with weak enforcement of boundary conditions leads to improved solution accuracies. Furthermore, we show that a feasible solution exists to the approximation of the constrained nonlinear optimal control problem. Using an example, the Galerkin method described in this paper is shown to be more accurate than some existing methods.

**Index Terms**—Constrained optimal control, Galerkin, pseudospectral.

## I. INTRODUCTION

SINCE the mid 1990's, pseudospectral (PS) methods, have gained popularity for solving optimal control problems [2], [7], [8], [9], [10], [13], [14]. PS methods have produced accurate solutions on a wide-range of optimal control problems. Recent highlights include the use of the Legendre PS method (LPM) for attitude control of the International Space Station (ISS) [14] and the first ever minimum-time rotational maneuver performed in orbit by a NASA space telescope called TRACE [2]. In the LPM, the problem state-control constraints are discretized using Legendre-Gauss-Lobatto (LGL) points. The states and controls are approximated with globally interpolating Lagrange polynomials and the cost function is approximated using LGL quadrature rule. Additionally with the LPM, problem end conditions are typically enforced with inequality endpoint constraints.

While the LPM has shown to be a good all-round method for solving nonlinear optimal control problems, enforcing boundary conditions (BCs) by means of separate constraints may introduce errors into the approximation. Alternatively, discretizing the dynamics using Galerkin techniques allows for the weak imposition of BCs. That is, end conditions may be enforced only up to the accuracy of the approximation itself. Galerkin formulations with weak enforcement of BCs have been shown to produce improved accuracies in many applications. A detailed discussion is given by Canuto et al. (see Section 3.7 of [6]). The Galerkin formulation with weak imposition of end conditions may also allow for problem discretizations with other than LGL grids, such as Legendre-Gauss (LG) and Legendre-Gauss-Radau (LGR) points. This

may be advantageous due to the increased accuracy of LG and LGR quadrature. A new direct method for solving optimal control problems is outlined in Section II which uses Galerkin techniques to discretize system dynamics.

Note that throughout this paper, we make use of the Sobolev spaces,  $W^{m,p}$ , that consist of all functions,  $u : [-1, 1] \rightarrow \mathbb{R}^n$  having weak derivative,  $u^{(i)} \in L^p$ , where  $0 \leq i \leq m$ , with the norm

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

and  $\|u\|_{L^p}$  defined by

$$\|u\|_{L^p} = \begin{cases} \left( \int_{-1}^1 |u(t)|^p dt \right)^{1/p}, & 1 \leq p < \infty, \\ \text{ess sup}_{t \in (-1,1)} |u(t)|, & p = \infty. \end{cases}$$

Also, we simplify notation by letting  $W^{m,2} = H^m$  and use  $\|u\|_{\infty}$  to denote the maximum element of vector,  $u \in \mathbb{R}^n$ .

## II. GALERKIN OPTIMAL CONTROL METHOD

In the following we introduce the Galerkin optimal control method (GOCM) for nonlinear problems. However, consider first the following constrained optimal control problem.

### A. Problem Statement

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

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

subject to the dynamics  $\dot{x}(t) = f(x(t), u(t))$ , endpoint conditions  $e(x(-1), x(1)) = 0$ , and mixed state-control path conditions  $h(x(t), u(t)) \leq 0$ . It is assumed that  $\mathcal{F} : \mathbb{R}^{N_x} \times \mathbb{R}^{N_u} \rightarrow \mathbb{R}$ ,  $\mathcal{E} : \mathbb{R}^{N_x} \times \mathbb{R}^{N_x} \rightarrow \mathbb{R}$ ,  $f : \mathbb{R}^{N_x} \times \mathbb{R}^{N_u} \rightarrow \mathbb{R}^{N_x}$ ,  $e : \mathbb{R}^{N_x} \times \mathbb{R}^{N_x} \rightarrow \mathbb{R}^{N_e}$ ,  $h : \mathbb{R}^{N_x} \times \mathbb{R}^{N_u} \rightarrow \mathbb{R}^{N_h}$  are Lipschitz continuous with respect to their argument.

### B. Numerical Solution Approach

There are four parts to the numerical solution to the optimal control problem using the direct method: discretization of the system dynamics, discretization of the state-control constraints, integration of the cost function and solving the nonlinear program (NLP). In the GOCM approach, we use a Galerkin method to discretize the system dynamics based on LGL quadrature points. The LGL points,  $\{t_k\}_{k=0}^N$ , are defined by  $t_0 = -1 < t_1 < \dots < t_N = 1$  and are the roots of  $(1-t^2)\dot{L}_N(t)$ , where  $\dot{L}_N(t)$  is the derivative of the  $N$ -th order Legendre polynomial,  $L_N(t)$ . The discretization

R. Boucher, Department of Applied Mathematics, Naval Postgraduate School, Monterey, CA 93940, rboucher@nps.edu

W. Kang, Department of Applied Mathematics, Naval Postgraduate School, Monterey, CA 93940, wkang@nps.edu

Q. Gong, Department of Applied Mathematics and Statistics, University of California, Santa Cruz, CA 95064, qigong@soe.ucsc.edu

works in the interval of  $[-1, 1]$  and will then provide the framework for our problem. LGL quadrature rule will be used to integrate the cost function. Finally, the resulting NLP can be solved using existing NLP algorithms.

### C. Method for Approximation

In a standard nodal Galerkin method, a solution to the differential equation  $\dot{x} - f(x(t), u(t)) = 0$  may be approximated with the weak integral form

$$\int_{-1}^1 \phi_i(t) (\dot{x}^N(t) - f(x^N(t), u^N(t))) dt = 0,$$

for  $i = 0, 1, \dots, N$ . The state and control functions,  $x(t)$  and  $u(t)$ , are approximated by  $N$ -th order Lagrange polynomials  $x(t) \approx x^N(t) = \sum_{i=0}^N \bar{x}_i \phi_i(t)$  and  $u(t) \approx u^N(t) = \sum_{i=0}^N \bar{u}_i \phi_i(t)$ , with  $\bar{x}_i \approx x(t_i)$  and  $\bar{u}_i \approx u(t_i)$ , for  $i = 0, 1, \dots, N$ . The test and basis functions,  $\{\phi_i\}_{i=0}^N$ , are Lagrange polynomials of order  $N$ , obtained from the general definition

$$\phi_i(t) = \prod_{\substack{j=0 \\ j \neq i}}^N \frac{(t-t_j)}{(t_i-t_j)},$$

and defined on a grid of LGL points,  $\{t_i\}_{i=0}^N$ . Integration by parts on the first term of the weak integral equation results in the Galerkin weak form

$$-\int_{-1}^1 \dot{\phi}_i x^N dt + [\phi_i x^N]_{-1}^1 - \int_{-1}^1 \phi_i f(x^N, u^N) dt = 0.$$

In terms of the approximating polynomials and true boundary conditions (letting  $x^N(-1) \rightarrow x(-1)$  and  $x^N(1) \rightarrow x(1)$ ) we have

$$\begin{aligned} & -\sum_{j=0}^N \int_{-1}^1 \dot{\phi}_i \phi_j dt \bar{x}_j - \phi_i(-1)x(-1) \\ & + \phi_i(1)x(1) - \int_{-1}^1 \phi_i f(x^N, u^N) dt = 0, \end{aligned}$$

for  $i = 0, 1, \dots, N$ , with the derivative of the Lagrange polynomial given by the relationship

$$\dot{\phi}_i(t) = \sum_{\substack{k=0 \\ k \neq i}}^N \left( \frac{1}{t_i - t_j} \right) \prod_{\substack{j=0 \\ j \neq i \\ j \neq k}}^N \frac{(t-t_j)}{(t_i-t_j)}.$$

Integration by parts, yet again, results in the Galerkin strong form with weak enforcement of BCs

$$\begin{aligned} & \sum_{j=0}^N D_{ij} \bar{x}_j + \phi_i(-1) \left( \sum_{j=0}^N \phi_j(-1) \bar{x}_j - x(-1) \right) \\ & - \phi_i(1) \left( \sum_{j=0}^N \phi_j(1) \bar{x}_j - x(1) \right) - c_i = 0. \end{aligned}$$

The  $(N+1) \times (N+1)$  Galerkin differentiation matrix,  $D$ , is defined by  $D_{ij} = \int_{-1}^1 \phi_i \dot{\phi}_j dt$ , for  $i, j = 0, 1, \dots, N$ , and can be calculated exactly with LGL quadrature (since LGL quadrature rule will provide zero error for polynomial integrands of less than or equal to  $2N-1$  [11]) by the relationship

$$D_{ij} = \sum_{k=0}^N \phi_i(t_k) \dot{\phi}_j(t_k) w_k, \quad i, j = 0, 1, \dots, N,$$

where the LGL weights,  $\{w_k\}_{k=0}^N$ , are defined by

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

Due to the property of the Lagrange polynomials

$$\phi_j(t_i) = \delta_{ij} = \begin{cases} 1, & i = j, \\ 0, & i \neq j, \end{cases}$$

the Galerkin differentiation matrix,  $D$ , is related to the Legendre PS differentiation matrix,  $A$ , by the simplified expression,  $D_{ij} = \dot{\phi}_j(t_i) w_i = A_{ij} w_i$ , where  $A_{ij} = \dot{\phi}_j(t_i)$  for  $i, j = 0, 1, \dots, N$ . The  $(N+1) \times 1$  right-hand-side (RHS) vector,  $c$ , is defined as  $c_i = \int_{-1}^1 \phi_i f(x^N, u^N) dt$ , for  $i = 0, 1, \dots, N$ , and can be approximated by the relationship

$$c_i \approx \sum_{k=0}^Q \phi_i(t_k) f \left( \sum_{j=0}^N \bar{x}_j \phi_j(t_k), \sum_{j=0}^N \bar{u}_j \phi_j(t_k) \right) w_k,$$

for  $i = 0, 1, \dots, N$ . Note that for LGL quadrature rule,  $Q = N+1$  integration points will integrate the RHS vector exactly when  $f(x(t), u(t))$  is linear in  $x(t)$  and  $u(t)$ . In the case of a nonlinear function  $f$ , the accuracy of integration may also be improved by increasing the number of quadrature points,  $Q$ . However, increasing  $Q$  may significantly add to computation time due to the required interpolation of the state and control vectors. If  $Q = N$  LGL quadrature nodes are used, the approximation to the RHS vector simplifies to  $c_i \approx \bar{c}_i = f(\bar{x}_i, \bar{u}_i) w_i$ , for  $i = 0, 1, \dots, N$ . Therefore the system may be approximated by the relationship

$$\sum_{j=0}^N D_{ij} \bar{x}_j + \kappa_i - \bar{c}_i = 0,$$

for  $i = 0, 1, \dots, N$ , where

$$\kappa_i = \begin{cases} \bar{x}_0 - x(-1), & i = 0, \\ x(1) - \bar{x}_N, & i = N, \\ 0, & i \neq 0, N. \end{cases}$$

**Remark 1** The BC vector,  $\kappa$ , now provides a natural way to introduce end conditions into our numerical scheme. BCs such as  $e(x(-1), x(1)) = [x(-1) - \alpha, x(1) - \beta]^T = [0, 0]^T$  may be imposed by defining  $\kappa$  as

$$\kappa_i = \begin{cases} \bar{x}_0 - \alpha, & i = 0, \\ \beta - \bar{x}_N, & i = N, \\ 0, & i \neq 0, N, \end{cases}$$

for  $i = 0, 1, \dots, N$ . This allows for much freedom in the enforcement of BCs. With the GOCM with weak enforcement of end conditions, BCs can now be enforced in a weak sense. In other words, end conditions can be imposed only up to the order of accuracy of the numerical approximation itself, which is sufficient for many applications. In the case of a problem with an incomplete set of BCs, such as an initial value problem with condition  $e(x(-1), x(1)) = x(-1) - \alpha = 0$ ,  $\kappa$  may be defined as

$$\kappa_i = \begin{cases} \bar{x}_0 - \alpha, & i = 0, \\ 0, & i = N, \\ 0, & i \neq 0, N. \end{cases}$$

Lastly, for more complex BCs such as periodic conditions  $e(x(-1), x(1)) = x(-1) - x(1) = 0$ , or other complicated BCs such as nonlinear functions of  $x(-1)$  and  $x(1)$ ,  $\kappa$  may be defined as  $\kappa_i = 0$ , for  $i = 0, 1, \dots, N$ , and the condition  $e(x(-1), x(1)) = 0$  may be enforced as a set of constraints by the NLP. However, this last case will result in strong enforcement of the BCs and some of the advantages of using the GOCM may be lost.

**Remark 2** The weak integral formulation also allows for consideration of quadrature points that don't include both endpoints, such as LG and LGR, or flipped-LGR (F-LGR) nodes. This is advantageous since LG and LGR quadrature rules may lead to increased accuracies when performing the RHS vector integration. LG quadrature rule integration is exact for polynomial integrands of degree less than or equal to  $2N + 1$ , where the LG nodes,  $\{t_k\}_{k=0}^N$ , are defined by  $-1 < t_0 < \dots < t_N < 1$ , and are the roots of  $L_{N+1}(t)$ . LGR quadrature rule is exact for polynomial integrands of degree less than or equal to  $2N$ , where the LGR nodes,  $\{t_k\}_{k=0}^N$ , are defined by  $t_0 = -1 < t_1 < \dots < t_N < 1$ , and are the roots of  $L_{N+1}(t) + L_N(t)$ . Lastly, the F-LGR nodes are the negative of the LGR points. These details and others, with regard to node selection (such as quadrature weights), are given by Canuto et al. [6].

**Remark 3** Due to the nilpotent (and therefore invertible) nature of the Galerkin differentiation matrix [12],  $D$ , a feasible solution to the equality dynamical constraint may not exist. In order to guarantee feasibility of the discretized problem, the following inequality constraint is suggested,

$$\left\| \sum_{j=0}^N D_{ij} \bar{x}_j + \kappa_i - \bar{c}_i \right\|_{\infty} \leq \delta^N, \quad i = 0, 1, \dots, N,$$

where  $\delta^N$  is the feasibility tolerance which is dependent on  $N$  and the smoothness of  $x$  and  $u$ .

#### D. Computation Strategy

The computation strategy for the GOCM with weak enforcement of BCs is presented in two forms. First, the strategy for the continuous problem, in terms of the approximating polynomials is outlined, denoted as GOCM-W. Next, the discrete problem, discretized on a LGL grid is presented, denoted as GOCM-W. Note that the computational strategies as well as feasibility theorems (presented in Section III) consider the following two cases:

- Case 1:  $u(\cdot)$  piecewise  $C^0$  &  $x(\cdot) \in C^0$  & piecewise  $C^1$ ,
- Case 2:  $u(\cdot) \in H^{m-1}$  and  $x(\cdot) \in H^m$  with  $m \geq 2$ .

**GOCM-W:** Find the feasible solution  $(x^N(t), u^N(t))$  that minimizes

$$\mathcal{J}(x^N(t), u^N(t))$$

$$= \int_{-1}^1 \mathcal{F}(x^N(t), u^N(t)) dt + \mathcal{E}(x^N(-1), x^N(1)),$$

subject to the Galerkin constraints

$$\begin{aligned} & \left\| \int_{-1}^1 \phi_k(t) (\dot{x}^N(t) - f(x^N(t), u^N(t))) - \kappa_k \right\|_{\infty} \\ & \leq MN^{-\alpha}, \quad k = 0, 1, \dots, N \end{aligned}$$

and  $\|h^+(x^N(t), u^N(t))\|_{L^2} \leq MN^{-\alpha}$ , where  $\alpha = 1/2$  and  $(m-1)$ , for Case 1 and 2, respectively;  $M$  is a positive constant, independent of  $N$ ; and

$$\kappa_k = \begin{cases} x^N(-1) - \alpha, & k = 0, \\ \beta - x^N(1), & k = N, \\ 0, & k \neq 0, N, \end{cases} \quad h^+ = \begin{cases} h, & h > 0, \\ 0, & h \leq 0, \end{cases}$$

where  $e(x(-1), x(1)) = [x(-1) - \alpha, x(1) - \beta]^T = [0, 0]^T$ .

**GOCM-W:** Find the feasible solution  $(\bar{X}, \bar{U})$  that minimizes  $\mathcal{J}(\bar{X}, \bar{U}) = \sum_{k=0}^N \mathcal{F}(\bar{x}_k, \bar{u}_k) w_k + \mathcal{E}(\bar{x}_0, \bar{x}_N)$ , where  $\bar{X} = [\bar{x}_0, \dots, \bar{x}_N]$  and  $\bar{U} = [\bar{u}_0, \dots, \bar{u}_N]$ , subject to the Galerkin constraints  $\left\| \sum_{j=0}^N D_{kj} \bar{x}_j + \kappa_k - \bar{c}_k \right\|_{\infty} \leq MN^{-\alpha}$  and  $h(\bar{x}_k, \bar{u}_k) \leq MN^{-\alpha} \cdot \mathbf{1}$ , for  $k = 0, 1, \dots, N$ , where  $\alpha = 1/2$  and  $(m-1)$ , for Case 1 and 2, respectively;  $\mathbf{1}$  denotes  $[1, \dots, 1]^T$ ;  $M$  is a constant, independent of  $N$ ; and

$$\kappa_k = \begin{cases} \bar{x}_0 - \alpha, & k = 0, \\ \beta - \bar{x}_N, & k = N, \\ 0, & k \neq 0, N, \end{cases}$$

where  $e(x(-1), x(1)) = [x(-1) - \alpha, x(1) - \beta]^T = [0, 0]^T$ .

**Remark 4** The optimization problem can be solved with a commercial sequential quadratic programming software package. A feasible solution can be found that satisfies the tolerances specified in the NLP by adjusting the order of polynomial used in the approximation. A number of modifications may be made to the algorithm to improve the accuracy of the numerical solution. These options include: changing node type (from LGL to LG or LGR), over-integration of the RHS term, and changing the way that BCs are enforced. In the case of boundary condition enforcement, if strong enforcement of the end conditions is desired, the following additional constraint may be added to the optimization Problems GOCM-W and GOCM-W:  $\|e(x^N(-1), x^N(1))\|_{\infty} \leq MN^{-\alpha}$ . The algorithm can be modified to accommodate constraints on one or more of the end conditions as required. The example in section IV will compare a number of the options available. For the remainder of this paper, the abbreviation GOCM-S will represent the GOCM with strong enforcement of the end conditions.

### III. FEASIBILITY OF GOCM-W AND GOCM-W

In order to guarantee feasibility of GOCM-W and GOCM-W, Theorems 1 and 2 show that a relaxation of the dynamical equality constraint to inequality is required, but first consider the following three lemmas.

**Lemma 1** [6] Let  $p^N(t)$  be the  $N$ -th order truncated Legendre series approximation to  $\xi \in H^m$ ,  $t \in [-1, 1]$ , then

$$\|\xi(t) - p^N(t)\|_{L^2} \leq C_1 C_0 N^{-m},$$

$$\|\xi(t) - p^N(t)\|_{L^\infty} \leq C_3 C_2 N^{1/2-m},$$

$\forall t \in [-1, 1]$ , where  $C_1$  and  $C_3$  are constants, independent of  $N$ ,  $C_0 = \|\xi\|_{H^m}$ ,  $C_2 = V(\xi^{(m)})$ ,  $V(\xi^{(m)})$  is the total variation of  $\xi^{(m)}$ , and  $m \geq 0$ . Note that  $p^N(t)$  with the smallest norm  $\|\xi(t) - p^N(t)\|_{L^2}$  is called the  $N$ -th order best polynomial approximation of  $\xi$  in the  $L^2$ -norm.

**Proof:** This is a standard result in approx. theory; see [6].

**Lemma 2** Let  $\zeta(t) = g(t) + hu_{t_c}(t)$ ,  $t \in [-1, 1]$ , where  $\zeta, u_{t_c} \in H^0$ ,  $g \in H^1$ , and  $u_{t_c}(t) = u(t - t_c)$  is the unit step function defined by

$$u_{t_c}(t) = \begin{cases} 0, & -1 \leq t < t_c, \\ 1, & t_c \leq t \leq 1. \end{cases}$$

Also, let  $p^N(t) = \sum_{n=0}^N p_n L_n$  be the  $N$ -th order truncated Legendre series polynomial approximation of  $\zeta$ . Then

$$\|\zeta(t) - p^N(t)\|_{L^2} \leq C_1 C_0 N^{-1} + C_2 N^{-1/2},$$

$\forall t \in [-1, 1]$ ,  $t_c \neq -1, 1$  and  $|h| < \infty$ , where  $C_1$  and  $C_2(t_c, h)$  are constants, independent of  $N$ , and  $C_0 = \|g\|_{H^1}$ .

**Proof:** Proof is omitted due to conf. space limits; see [5].

**Lemma 3** Let  $\{\phi_i\}_{i=0}^N$  be the Lagrange polynomials of order  $N$  defined on LGL grid,  $\{t_i\}_{i=0}^N$ . Then  $\|\phi_i\|_{L^2} \leq C_0 w_i^{1/2}$ , for each  $i = 0, 1, \dots, N$ , where  $w_i$  is the LGL weight at gridpoint,  $t_i$ , and  $C_0$  is a positive constant independent of  $N$ .

**Proof:** From [6], the discrete norm,  $\|\xi_N\|_N = \left( \sum_{k=0}^N |\xi_k|^2 w_k \right)^{1/2}$ , has the property  $\|\xi_N\|_{L^2} \leq C_0 \|\xi_N\|_N$  for  $\xi_N \in P_N$ , where  $C_0$  is a positive constant, independent of  $N$ . Since  $\phi_i \in P_N$ , for all  $i = 0, 1, \dots, N$ , we have  $\|\phi_i\|_{L^2} \leq C_0 \|\phi_i\|_N$ . Further, from the property of the Lagrange polynomial,  $\phi_j(t_i) = \delta_{ij}$ , we have  $\|\phi_i\|_N = \left( \sum_{k=0}^N |\phi_i(t_k)|^2 w_k \right)^{1/2} = w_i^{1/2}$ , where  $\{w_i\}_{i=0}^N$  are the LGL quadrature weights associated with the LGL points,  $\{t_i\}_{i=0}^N$ . Hence  $\|\phi_i\|_{L^2} \leq C_0 w_i^{1/2}$ .

**Theorem 1** Given any feasible solution  $t \mapsto (x, u)$ , for Problem B, consider the following two cases:

Case 1:  $u(\cdot)$  piecewise  $C^0$  &  $x(\cdot) \in C^0$  & piecewise  $C^1$ ,

Case 2:  $u(\cdot) \in H^{m-1}$  and  $x(\cdot) \in H^m$  with  $m \geq 2$ .

Then, there exists a positive integer  $N_0$  such that, for any  $N \geq N_0$ , GOCM- $\bar{W}$  has a feasible solution,  $(x^N(t), u^N(t))$ , such that  $\|x(t) - x^N(t)\|_{L^2} \leq MN^{-\alpha}$  and  $\|u(t) - u^N(t)\|_{L^2} \leq MN^{-\alpha}$ , where  $\alpha = 1/2$  and  $(m-1)$ , for Case 1 and 2, respectively;  $M$  is a positive constant, independent of  $N$ .

**Proof:** Let  $p(t)$  be the  $(N-1)$ -th order best polynomial approximation of  $\dot{x}(t)$  in the  $L^2$ -norm. By Lemmas 1 and 2 there is a constant  $C_1$  independent of  $N$ , for any  $N \geq N_1$ , such that  $\|\dot{x}(t) - p(t)\|_{L^2} \leq C_1 N^{-\alpha}$ , where  $\alpha = 1/2$  and  $(m-1)$ , for Case 1 and 2, respectively. Define  $x^N(t) = \int_{-1}^t p(s) ds + x(-1)$ . Then  $p(t) = \dot{x}(t)$  and  $\|x(t) - x^N(t)\|_{L^2} \leq 2C_1 N^{-\alpha}$ , since, from Hölder's inequality [1], we have

$$\begin{aligned} |x(t) - x^N(t)| &= \left| \int_{-1}^t (\dot{x}(s) - p(s)) ds \right| \\ &\leq \sqrt{2} \|\dot{x}(t) - p(t)\|_{L^2} \leq \sqrt{2} C_1 N^{-\alpha}. \end{aligned}$$

Also, let  $u^N(t)$  be the  $N$ -th order Legendre polynomial so that  $\|u(t) - u^N(t)\|_{L^2} \leq C_2 N^{-\alpha}$ . From our Galerkin approximation (recall that the BC term,  $\kappa_k = 0$ , for  $k \neq 0, N$ ), Hölder's inequality, and Lemma 3, we have for each  $k = 1, 2, \dots, N-1$ ,

$$\begin{aligned} & \left| \int_{-1}^1 \phi_k(t) (\dot{x}^N(t) - f(x^N(t), u^N(t))) dt \right| \\ & \leq \int_{-1}^1 |\phi_k(t)| |\dot{x}^N(t) - f(x^N(t), u^N(t))| dt \\ & \leq \|\phi_k(t)\|_{L^2} \|\dot{x}^N(t) - f(x^N(t), u^N(t))\|_{L^2} \\ & \leq C_3 w_k^{1/2} \|\dot{x}^N(t) - f(x^N(t), u^N(t))\|_{L^2} \\ & \leq C_3 w_k^{1/2} \|\dot{x}(t) - \dot{x}^N(t)\|_{L^2} \\ & \quad + C_3 w_k^{1/2} \|\dot{x}(t) - f(x^N(t), u^N(t))\|_{L^2} \\ & = C_3 w_k^{1/2} \|\dot{x}(t) - p(t)\|_{L^2} \\ & \quad + C_3 w_k^{1/2} \|f(x(t), u(t)) - f(x^N(t), u^N(t))\|_{L^2} \\ & \leq C_1 C_3 w_k^{1/2} N^{-\alpha} + C_3 L_1 w_k^{1/2} \|x(t) - x^N(t)\|_{L^2} \\ & \quad + C_3 L_2 w_k^{1/2} \|u(t) - u^N(t)\|_{L^2} \\ & \leq C_3 (C_1 + 2C_1 L_1 + C_2 L_2) w_k^{1/2} N^{-\alpha} = C_4 w_k^{1/2} N^{-\alpha}, \end{aligned}$$

where  $C_4 = C_3 (C_1 + 2C_1 L_1 + C_2 L_2)$  and  $L_1$  and  $L_2$  are the Lipschitz constants of  $f$  with respect to  $x$  and  $u$ , respectively, which are all independent of  $N$ . For  $k = 0$  we have

$$\begin{aligned} & \left| \int_{-1}^1 \phi_0(t) (\dot{x}(t) - f(x^N(t), u^N(t))) dt + \kappa_0 \right| \\ & \leq C_4 w_0^{1/2} N^{-\alpha} + |x^N(-1) - x(-1)| = C_4 w_0^{1/2} N^{-\alpha}, \end{aligned}$$

since  $|x^N(-1) - x(-1)| = 0$ . For  $k = N$ , we have

$$\begin{aligned} & \left| \int_{-1}^1 \phi_N(t) (\dot{x}(t) - f(x^N(t), u^N(t))) dt + \kappa_N \right| \\ & \leq C_4 w_N^{1/2} N^{-\alpha} + |x(1) - x^N(1)| \\ & \leq C_4 w_N^{1/2} N^{-\alpha} + \sqrt{2} C_1 N^{-\alpha}. \end{aligned}$$

Finally, for each  $k = 0, 1, \dots, N$ , we have  $\left| \int_{-1}^1 \phi_k(t) (\dot{x}(t) - f(x^N(t), u^N(t))) dt + \kappa_k \right| \leq C_5 N^{-\alpha}$ , for all  $N > N_2$ , where  $C_5$  is a constant, independent of  $N$ .

For the path constraint,  $h$ , let  $D = \{t | h^N > 0\}$ ,  $\bar{D} = [-1, 1] \setminus D$ , where  $h^N = h(x^N(t), u^N(t))$ . Then, since  $h \leq 0$ , we have

$$\begin{aligned}
& \|h^+(x^N(t), u^N(t))\|_{L^2} = \left( \int_D (h^N)^2 dt \right)^{1/2} \\
& \leq \left( \int_D (h^N - h)^2 dt \right)^{1/2} \\
& \leq \left( \int_D (h^N - h)^2 dt + \int_{\bar{D}} (h^N - h)^2 dt \right)^{1/2} \\
& = \left( \int_{-1}^1 \int_D (h^N - h)^2 dt \right)^{1/2} = \|h^N - h\|_{L^2} \\
& \leq L_3 \|x(t) - x^N(t)\|_{L^2} + L_4 \|u(t) - u^N(t)\|_{L^2} \\
& \leq (2C_1 L_3 + C_2 L_4) N^{-\alpha},
\end{aligned}$$

where  $L_3$  and  $L_4$  are the Lipschitz constants of  $h$  with respect to  $x$  and  $u$ , respectively, which are independent of  $N$ . Thus a solution  $(x^N(t), u^N(t))$  to GOCM-W is feasible!

**Theorem 2** Given any feasible solution  $t \mapsto (x, u)$ , for Problem B, consider the following two cases:

Case 1:  $u(\cdot)$  piecewise  $C^0$  &  $x(\cdot) \in C^0$  & piecewise  $C^1$ ,

Case 2:  $u(\cdot) \in H^{m-1}$  and  $x(\cdot) \in H^m$  with  $m \geq 2$ .

Then, there exists a positive integer  $N_0$  such that, for any  $N \geq N_0$ , Problem GOCM-W has a feasible solution,  $(\bar{X}, \bar{U})$ , such that  $\|x(t) - x^N(t)\|_{L^2} \leq MN^{-\alpha}$ , where  $\alpha = 1/2$  and  $(m-1)$ , for Case 1 and 2, respectively;  $M$  is a positive constant, independent of  $N$ .

**Proof:** Let  $p(t)$  be the  $(N-1)$ -th order polynomial approximation of  $\dot{x}(t)$  in the  $L^2$ -norm. By Lemma 1 there is a constant  $C_1$  independent of  $N$ , for any  $N \geq N_1$ , such that  $\|\dot{x}(t) - p(t)\|_{L^\infty} \leq C_1 N^{-\beta}$ , where  $\beta = -1/2$  and  $(m-3/2)$ , for Case 1 and 2, respectively. Also, by Lemmas 1 and 2 there is a constant  $C_2$  independent of  $N$ , for any  $N \geq N_2$ , such that  $\|\dot{x}(t) - p(t)\|_{L^2} \leq C_2 N^{-\alpha}$ , where  $\alpha = 1/2$  and  $(m-1)$ , for Case 1 and 2, respectively. Define  $x^N(t) = \int_{-1}^t p(s) ds + x(-1)$ . Then  $p(t) = \dot{x}(t)$  and  $\|x(t) - x^N(t)\|_{L^2} \leq 2C_2 N^{-\alpha}$ , since, from Hölder's inequality, we have

$$\begin{aligned}
|x(t) - x^N(t)| &= \left| \int_{-1}^t (\dot{x}(s) - p(s)) ds \right| \\
&\leq \sqrt{2} \|\dot{x}(t) - p(t)\|_{L^2} \leq \sqrt{2} C_2 N^{-\alpha}.
\end{aligned}$$

Also, let  $u^N(t)$  be an interpolating function of  $u(t)$  defined as  $u^N(t) = \sum_{j=0}^N \bar{u}_j \psi_j(t)$ , where  $\bar{u}_j = u(t_j)$ , and  $\psi(t)$  is any continuous function (not necessarily a polynomial) with the cardinality property  $\psi_j(t_i) = \delta_{ij}$ . Since  $x^N(t)$  is a  $N$ -th order polynomial, we have [6]  $\dot{x}^N(t_k) = \sum_{j=0}^N A_{kj} \bar{x}_j$ , where  $A$  is the  $(N+1) \times (N+1)$  Legendre PS differentiation matrix and  $\bar{x}_j = x^N(t_j)$ . From [6], the LGL weights,  $\{w_k\}_{k=0}^N$ , have the property  $w_k \leq C_3 N^{-1} (1 - (t_k)^2)^{1/2}$ , for  $k = 1, 2, \dots, N-1$  and constant  $C_3 > 0$  independent of  $k$  and  $N$ . For  $k = 0, N$ , we have  $w_0, w_N = \frac{2}{N(N+1)}$ . From our Galerkin approximation, for each  $k = 1, \dots, N-1$  (recall that the BC term,  $\kappa_k = 0$ , for  $k \neq 0, N$ ), we have

$$\left| \sum_{j=0}^N D_{kj} \bar{x}_j - \bar{c}_k \right| = \left| \sum_{j=0}^N A_{kj} \bar{x}_j - f(\bar{x}_k, \bar{u}_k) \right| w_k$$

$$\begin{aligned}
& = |\dot{x}^N(t_k) - f(\bar{x}_k, \bar{u}_k)| w_k \\
& \leq |p(t_k) - f(x(t_k), u(t_k))| w_k \\
& + |f(x(t_k), u(t_k)) - f(\bar{x}_k, \bar{u}_k)| w_k \\
& = |p(t_k) - \dot{x}(t_k)| w_k \\
& + |f(x(t_k), u(t_k)) - f(\bar{x}_k, \bar{u}_k)| w_k \\
& \leq \|p(t) - \dot{x}(t)\|_{L^\infty} w_k + L_1 |x(t_k) - x^N(t_k)| w_k \\
& \leq C_1 w_k N^{-\beta} + \sqrt{2} L_1 C_2 w_k N^{-\alpha},
\end{aligned}$$

where  $L_1$  is the Lipschitz constant of  $f$  with respect to  $x$ , which is independent of  $N$ . For  $k = 0$  we have

$$\begin{aligned}
\left| \sum_{j=0}^N D_{0j} \bar{x}_j + \kappa_0 - \bar{c}_0 \right| &\leq \left| \sum_{j=0}^N D_{0j} \bar{x}_j - \bar{c}_0 \right| + |\bar{x}_0 - x(-1)| \\
&\leq C_1 w_0 N^{-\beta} + \sqrt{2} L_1 C_2 w_0 N^{-\alpha},
\end{aligned}$$

since  $|x^N(-1) - x(-1)| = 0$ . For  $k = N$ , we have

$$\begin{aligned}
\left| \sum_{j=0}^N D_{Nj} \bar{x}_j + \kappa_N - \bar{c}_N \right| &\leq \left| \sum_{j=0}^N D_{Nj} \bar{x}_j - \bar{c}_N \right| + |x(1) - \bar{x}_N| \\
&= C_1 w_N N^{-\beta} + \sqrt{2} C_2 (L_1 w_N + 1) N^{-\alpha}.
\end{aligned}$$

Finally, for each  $k = 0, 1, \dots, N$ , since  $w_k \leq C_3 N^{-1}$ , we have  $\left| \sum_{j=0}^N D_{kj} \bar{x}_j + \kappa_k - \bar{c}_k \right| \leq C_4 N^{-\alpha}$ , for all  $N > N_3$ , where  $C_4$  is a constant, independent of  $N$ .

For the path constraint, the following estimate holds

$$\begin{aligned}
& |h(x(t), u(t)) - h(x^N(t), u^N(t))| \\
& \leq L_2 |x(t) - x^N(t)| \leq \sqrt{2} L_2 C_2 N^{-\alpha},
\end{aligned}$$

for each  $k = 0, 1, \dots, N$ , where  $L_2$  is the Lipschitz constant of  $h$  with respect to  $x$ . Hence

$$\begin{aligned}
h(\bar{x}_k, \bar{u}_k) &\leq h(x(t_k), u(t_k)) + \sqrt{2} L_2 C_2 N^{-\alpha} \cdot \mathbf{1} \\
&\leq \sqrt{2} L_2 C_2 N^{-\alpha} \cdot \mathbf{1}.
\end{aligned}$$

Thus a solution  $(\bar{X}, \bar{U})$  to GOCM-W is feasible!

#### IV. NONLINEAR INITIAL VALUE PROBLEM

Consider the two-dimensional initial value problem given by Gong et al. [9] of minimizing the cost function  $\mathcal{J} = 4x_1(2) + x_2(2) + 4 \int_0^2 u^2(t) dt$ , subject to the dynamics,  $\dot{x}_1(t) = x_2^3(t)$  and  $\dot{x}_2(t) = u(t)$ , and initial conditions,  $x_1(0) = 0$  and  $x_2(0) = 1$ . The analytic solution to this problem is given by

$$x_1(t) = \frac{2}{5} - \frac{64}{5(2+t)^5}, \quad x_2(t) = \frac{4}{(2+t)^2} \quad \text{and} \quad u(t) = \frac{-8}{(2+t)^3},$$

obtained via Pontryagin's maximum principle. This problem was solved using the GOCM-W with LGL nodes with optimality and feasibility tolerances of  $10^{-16}$  and  $10^{-16}$ , respectively. For an accuracy analysis of the GOCM, the exact solution was used as an initial guess. A demonstration of the robustness of the method can be found in [16]. Fig. 1 shows the results using  $n = 20$  nodes.

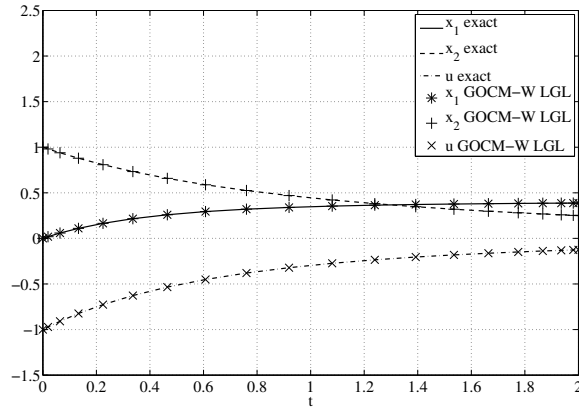


Fig. 1. Comparison of the exact and numerical solutions using the GOCM-W with  $n = 20$  LGL points.

The maximum errors from the true control,  $u$ , are shown in Fig. 2 for various approximations. Note that the GOCM-S numerical solution of  $u$  shows relatively the same exponential convergence rate as the LPM numerical solution. When using the GOCM-W with LGL and F-LGR nodes, the numerical solutions of  $u$  show a marked improvement over the GOCM-S, particularly with the lower order approximations. The GOCM-W approximation of  $u$  with  $n = 14$  LGL nodes achieves an impressive  $O(10^{-9})$  accuracy, similar to that of the GOCM-S numerical solution with  $n = 20$  nodes.

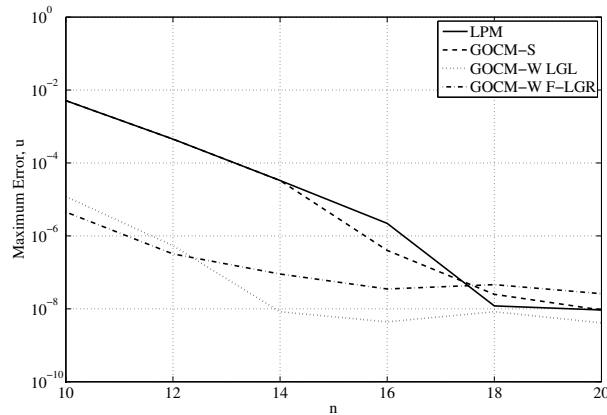


Fig. 2. Convergence of the numerical approximations of the control,  $u$ , for the LPM, GOCM-S and GOCM-W with LGL and F-LGR nodes.

## V. CONCLUSION

The Galerkin optimal control method introduced in this paper shows great potential in solving constrained nonlinear optimal control problems. A great strength of the GOCM is the ability to be modified to suit the needs of the optimal control problem at hand. From altering how boundary conditions are enforced, to changing the node type used, GOCM can potentially improve the accuracy of any LPM

solution [4]. In addition to the algorithm options mentioned here, Galerkin methods (and thus the GOCM) are easily extended to element-based methods, both continuous and discontinuous [3]. Hence, GOCM has the potential to tackle a wide variety of problems, to include those with discontinuous solutions and multi-stage problems. The GOCM also has the capability to be effectively applied to multi-scale problems, those in which the state and control variables may be approximated at different orders. Finally, it is important to note that there is an increasing body of work dealing with PDE control. Although not considered here, Galerkin methods are a standard choice for these types of problems.

## ACKNOWLEDGMENTS

This work was supported in part by the U.S. Air Force Office of Scientific Research (AFOSR) and the U.S. Naval Research Laboratory (NRL).

## REFERENCES

- [1] G. Bachman and L. Narici. *Functional Analysis*. New York, Academy Press, 1966.
- [2] N. Bedrossian, M. Karpenko and S. Bhatt, "Overclock My Satellite," *IEEE Spectrum*, 2012.
- [3] R. Boucher, W. Kang and Q. Gong, "Discontinuous Galerkin Optimal Control for Constrained Nonlinear Problems," *Proceedings of the International Conference on Control & Automation*, no. WeC3.5, Taichung, Taiwan, June 2014.
- [4] R. Boucher, W. Kang and Q. Gong, "Galerkin Optimal Control for Constrained Nonlinear Problems," *Proceedings of the American Control Conference*, no. ThA18.4, Portland, OR, June 2014.
- [5] R. Boucher, "Galerkin Optimal Control," tech. rep., Naval Postgraduate School, Monterey, CA, March 2014.
- [6] C. Canuto, M. Y. Hussaini, A. Quarteroni and T. A. Zang, *Spectral Methods, Fundamentals in Single Domains*, Berlin & Heidelberg, Germany: Springer-Verlag, 2006.
- [7] G. Elnagar, M. A. Kazemi and M. Razzaghi, "The Pseudospectral Legendre Method for Discretizing Optimal Control Problems," *IEEE Transactions on Automatic Control*, vol. 40, pp. 1793-1796, 1995.
- [8] F. Fahroo and I. M. Ross, "Costate Estimation by a Legendre Pseudospectral Method," *Proceedings of the AIAA Guidance, Navigation and Control Conference*, no. AIAA 1998-4222, Boston, MA, August 1998.
- [9] Q. Gong, W. Kang and I. M. Ross, "A Pseudospectral Method for the Optimal Control of Constrained Feedback Linearizable Systems," *IEEE Transactions on Automatic Control*, vol. 51, pp. 1115-1129, 2006.
- [10] Q. Gong, I. M. Ross, W. Kang and F. Fahroo, "Connections Between the Covector Mapping Theorem and the Convergence of Pseudospectral Methods for Optimal Control," *Computational Optimization and Applications*, vol. 41, pp. 307-335, 2008.
- [11] J. S. Hesthaven, S. Gottlieb and D. Gottlieb, *Spectral Methods for Time-Dependent Problems*, Cambridge: Cambridge University Press, 2007.
- [12] J. S. Hesthaven and T. Warburton, *Nodal Discontinuous Galerkin Methods, Algorithms, Analysis, and Application*, New York: Springer Science+Business Media, LLC, 2000.
- [13] W. Kang, "Rate of Convergence for a Legendre Pseudospectral Optimal Control of Feedback Linearizable Systems," *J. of Control Theory and Applications*, vol. 8, pp. 391-405, 2010.
- [14] W. Kang and N. Bedrossian, "Pseudospectral Optimal Control Theory Makes Debut Flight, Saves NASA \$1M in Under Three Hours," *SIAM News*, vol. 40, 2007.
- [15] G. Sansone, A. Diamond and E. Hille, *Orthogonal Functions*, New York: Robert E. Krieger Publishing Co., 1977.
- [16] M. Zhu, S. Lai, R. Boucher, B. Chen, X. Cheng and W. Kang, "Minimum Time Trajectory for Helicopter UAVs: Computation and Flight Test," *Applied Mathematical Sciences*, vol. 7, pp. 6475-6487, 2013.