2002

A Direct Method for Solving Nonsmooth Optimal Control Problems

Ross, I. Michael; Fahroo, Fariba

International Federation of Automatic Control

http://hdl.handle.net/10945/29652

15th Triennial World Congress, Barcelona, Spain

A DIRECT METHOD FOR SOLVING NONSMOOTH OPTIMAL CONTROL PROBLEMS

I. Michael Ross [1]

* *Department of Aeronautics and Astronautics, Code AA/Ro, Naval Postgraduate School, Monterey, CA 93943*

Fariba Fahroo [2]

** *Department of Mathematics, Code MA/Ff, Naval Postgraduate School, Monterey, CA 93943*

Abstract. We present a class of efficient direct methods for solving nonsmooth dynamic optimization problems where the dynamics are governed by controlled differential inclusions. Our methods are based on pseudospectral approximations of the differential constraints that are assumed to be given in the form of controlled differential inclusions including the usual vector field and differential-algebraic forms. Discontinuities in states, controls, cost functional, dynamic constraints and various other mappings associated with the generalized Bolza problem are allowed by the concept of pseudospectral knots which we introduce in this paper. The computational optimal control problem is reduced to a structured sparse nonlinear programming problem. A simple but illustrative moon-landing problem demonstrates our method.

Keywords. Optimal control, discontinuities, discretization, nonlinear programming.

1. INTRODUCTION

In this paper we introduce new methods for solving a broad *class of nonsmooth optimal control problems* defined over a finite time interval that may be fixed or free. We use the term nonsmooth in the sense of Clarke (Clarke, 1990) "... to refer to certain situations in which the smoothness (differentiability) of the data is not necessarily postulated." Although nonsmooth analysis applies to "severely" pathological problems, in the design of numerical methods there are some fundamental limitations. For example, Clarke's nonsmooth constructs allow functions to be extended-valued including functions defined over the extended real line, $\mathbb{R} \cup \{\infty\}$. However, *numerical methods are fundamentally limited to finite variables.* Hence our methods (as any and all numerical methods) apply to the subset of problems that are numerically realizable. In these problems all functions are piecewise smooth with locally nonsmooth behavior. Further, the number of locations of the nonsmoothness (e.g the number of discontinuities) are assumed to be finite. In addition to these nonsmooth characteristics, we pose a problem formulation that allows for implicit constraints given in the form of "pre-defined" segments. Such problem formulations frequently arise in the mission design of interplanetary spacecraft trajectories.

The basic idea behind the solution method is to approximate this problem by pseudospectral techniques to a mathematical programming problem and then solve it numerically. Pseudospectral methods have been used extensively in solving problems in fluid dynamics (Canuto *et al.*, 1988; Gottlieb *et al.*, 1984). Their applications to solving smooth optimal control problems is quite new (Elnagar *et al.*, 1995; Fahroo and Ross, 2001b). Pseudospectral approximations are based on expanding the underlying functions in terms of interpolating polynomials which interpolate these functions at some specially chosen nodes. These nodes are zeros of orthogonal polynomials (or their derivatives) such as Legendre polynomials (Legendre-Gauss points) or Chebyshev polynomials (Chebyshev points). What distinguishes these methods are the choice of these nodes, subsequently the definition of the Lagrange interpolating polynomials. These methods are quite efficient and more accurate than the traditional collocation methods (Betts, 1998) in solving smooth optimal control problems, but their use in solving nonsmooth problems can cause major difficulties. For example, even point constraints in smooth problem formulations cannot be handled by the "smooth" pseudospectral method because the location of the point may

[1] E-mail: imross@nps.navy.mil
[2] E-mail: ffahroo@nps.navy.mil

not be at the pre-allocated Gauss node. Adding more nodes for mesh refinements could lead to inefficiencies and ill-conditioning of the discretized problem. Further, some practical problems contain empirical models based on table-lookups which are often nonsmooth data. Also, jump discontinuities in states (such as those encountered in the trajectory optimization of multi-stage launch vehicles) cannot be handled by these methods. Even for problems with smooth data, the solution may be nonsmooth (Clarke, 1990). In this case, pseudospectral methods exhibit the well-known Gibbs phenomenon (Fornberg, 1998) resulting from the approximation of a nonsmooth function by a finite number of smooth functions. These difficulties are fundamentally due to the use of global orthogonal polynomials and nodal points which have a predetermined distribution. This distribution of nodal points yields optimal interpolation, but offers no choice in the selection of the points.

In this paper, we overcome all of the numerical difficulties mentioned above by introducing *Pseudospectral Knotting Methods*. We introduce the concepts of *hard* and *soft knots* for a method based on the Legendre pseudospectral method, but it is readily applicable for other pseudospectral methods as well. In between these knots, the problem is discretized at the Legendre-Gauss-Lobatto (LGL) nodes. The discretization of the dynamic constraints is achieved by a differentiation operator that naturally allows for one-sided stencils at the knots while the integral associated with the cost function is approximated by a Gauss quadrature. Information across the knots are passed in the form of event conditions.

2. A PROTOTYPE NONSMOOTH PROBLEM

We define a discrete event or simply an *event,* in terms of multifunctions defined over discrete points in phase space. In a numerical setting, these multifunctions are given in terms of inequalities and equalities at discrete points. Thus, in this context, the definition of events subsumes the notion of boundary conditions. For simplicity in presentation, a prototype problem is defined in terms of one interior event (i.e. one non-boundary point). It will be apparent later that our method easily extends to more than one interior event. The single interior point in our presentation defines two *segments* or *subarcs*: the first one stretching from τ_0 to τ_e and the second from τ_e to τ_f, where τ_e is the event point. The three events (one interior and two boundary points) can be defined succinctly in terms of event conditions

$$\mathbf{e}_l \le \mathbf{e}\big(\mathbf{x}_0, \mathbf{x}_e^-, \mathbf{x}_e^+, \mathbf{x}_f; \tau_0, \tau_e, \tau_f\big) \le \mathbf{e}_u \tag{2.1}$$

where we use the notation $\mathbf{x}_0 = \mathbf{x}(\tau_0), \mathbf{x}_f = \mathbf{x}(\tau_f)$, and

$$\mathbf{x}_e^- = \lim_{\epsilon\uparrow 0} \mathbf{x}(\tau_e + \epsilon),\ \ \mathbf{x}_e^+ = \lim_{\epsilon\downarrow 0} \mathbf{x}(\tau_e + \epsilon)$$

and $\mathbf{e} : \mathbb{R}^{N_x} \times \mathbb{R}^{N_x} \times \mathbb{R}^{N_x} \times \mathbb{R}^{N_x} \times \mathbb{R} \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}^{N_e}$ where N_x is the dimension of the state vector, and N_e is the dimension of the event vector which is continuously differentiable with respect to its variables. Note that this last condition on $\mathbf{e}$ does not imply continuous state histories. Thus, the problem is defined as follows: Determine the trajectory-control pair, $[\tau_0, \tau_f] \ni \tau \mapsto \{\mathbf{x}(\cdot) \in \mathbb{R}^{N_x}, \mathbf{u}(\cdot) \in \mathbb{R}^{N_u}\}$, (where N_u is the dimension of the control vector) and the optimal event time $\tau_e \in (\tau_0, \tau_f)$ that minimize the generalized Bolza cost functional,

$$\begin{aligned} J[\mathbf{x}(\cdot), \mathbf{u}(\cdot), \tau_0, \tau_e, \tau_f] = E(\mathbf{x}_0, \mathbf{x}_e^-, \mathbf{x}_e^+, \mathbf{x}_f; \tau_0, \tau_e, \tau_f) \\ + \int_{\tau_0}^{\tau_f} F(\mathbf{x}(\tau), \mathbf{u}(\tau), \tau) d\tau \end{aligned} \tag{2.2}$$

subject to differential constraints given in terms of *controlled differential inclusions*

$$\mathbf{f}_l \le \mathbf{f}(\dot{\mathbf{x}}(\tau), \mathbf{x}(\tau), \mathbf{u}(\tau), \tau) \le \mathbf{f}_u \tag{2.3}$$

event conditions,

$$\mathbf{e}_l \le \mathbf{e}\big(\mathbf{x}_0, \mathbf{x}_e^-, \mathbf{x}_e^+, \mathbf{x}_f; \tau_0, \tau_e, \tau_f\big) \le \mathbf{e}_u \tag{2.4}$$

and *mixed* state-control *path constraints,*

$$\mathbf{g}_l \le \mathbf{g}(\mathbf{x}(\tau), \mathbf{u}(\tau), \tau) \le \mathbf{g}_u \tag{2.5}$$

Note that even if $\mathbf{g}(\cdot)$ is smooth it automatically forces nonsmooth trajectories (Vinter, 2000). It is apparent from our formulation that we have replaced the notion of multifunctions used in nonsmooth analysis by relations: a fundamental necessity for any numerical approach based on using real numbers. In any case, in all the relations above, an equality constraint may be obtained by simply setting the lower bound equal to the upper bound. The event time τ_e may be fixed or free and is used to demarcate various types of possible nonsmoothness in the trajectory $\tau \mapsto \mathbf{x}(\tau)$ and the maps $F(\cdot), \mathbf{f}(\cdot)$, and $\mathbf{g}(\cdot)$ which are defined as $E : \mathbb{R}^{N_x} \times \mathbb{R}^{N_x} \times \mathbb{R}^{N_x} \times \mathbb{R}^{N_x} \times \mathbb{R} \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$, $F : \mathbb{R}^{N_x} \times \mathbb{R}^{N_u} \times \mathbb{R} \to \mathbb{R}$, $\mathbf{f} : \mathbb{R}^{N_x} \times \mathbb{R}^{N_x} \times \mathbb{R}^{N_u} \times \mathbb{R} \to \mathbb{R}^{N_x}$, $\mathbf{g} : \mathbb{R}^{N_x} \times \mathbb{R}^{N_u} \times \mathbb{R} \to \mathbb{R}^{N_g}$ where N_g is the dimension of the path constraint vector. These functions may be nonsmooth at the event time τ_e but are continuously differentiable in the open intervals $(\tau_0, \tau_e), (\tau_e, \tau_f)$. The controls are allowed to be piecewise continuous with discontinuities of the first kind at points other than τ_e. At τ_e, one or more controls maybe a Dirac delta function that causes a jump in the state variables. We use superscript 1 to denote the functions in the first interval, and superscript 2 for functions on the second interval. Thus, we have

$$F(\cdot) = \begin{cases} F^1(\mathbf{x}(\tau), \mathbf{u}(\tau), \tau) & \tau \in [\tau_0, \tau_e] \\ F^2(\mathbf{x}(\tau), \mathbf{u}(\tau), \tau) & \tau \in [\tau_e, \tau_f] \end{cases} \tag{2.6}$$

$$\mathbf{f}(\cdot) = \begin{cases} \mathbf{f}^1(\dot{\mathbf{x}}(\tau), \mathbf{x}(\tau), \mathbf{u}(\tau), \tau) & \tau \in [\tau_0, \tau_e] \\ \mathbf{f}^2(\dot{\mathbf{x}}(\tau), \mathbf{x}(\tau), \mathbf{u}(\tau), \tau) & \tau \in [\tau_e, \tau_f] \end{cases} \tag{2.7}$$

$$\mathbf{g}(\cdot) = \begin{cases} \mathbf{g}^1(\mathbf{x}(\tau), \mathbf{u}(\tau), \tau) & \tau \in [\tau_0, \tau_e] \\ \mathbf{g}^2(\mathbf{x}(\tau), \mathbf{u}(\tau), \tau) & \tau \in [\tau_e, \tau_f] \end{cases} \tag{2.8}$$

From these discussions, it is apparent that the prototype problem easily extends to more than one interior event. It is worth noting that various scenarios can happen both physically and in terms of modeling across an event. For example, over segment 1, the dynamics of the problem may be given in terms of a differential inclusion,

$$\mathbf{f}_l^1 \le \mathbf{f}^1(\dot{\mathbf{x}}(\tau), \mathbf{x}(\tau), \tau) \le \mathbf{f}_u^1 \tag{2.9}$$

while in segment 2, it may be defined in terms of a differential algebraic equation,

$$\mathbf{f}^2(\dot{\mathbf{x}}(\tau), \mathbf{x}(\tau), \mathbf{u}(\tau), \tau) = \mathbf{0} \tag{2.10}$$

(i.e. $\mathbf{f}_l^2 = \mathbf{f}_u^2 = \mathbf{0}$.)

3. PSEUDOSPECTRAL KNOTTING METHODS

To directly solve the Bolza problem posed in the previous section, two basic discretizations are needed: one for the integral associated with the cost function and another for the dynamic constraints. In the traditional collocation methods, the Bolza problem is typically converted to a Mayer problem thus reducing the approximation issue to just the discretization of the dynamics. In any case, in the spectral method, we approximate the integral by a sum (using quadrature over the LGL points) while the derivative is approximated by a discrete differential operator.

In the pseudospectral approximation of the optimal control problem, the LGL node points lie in the computational interval $[-1, 1]$. The time coordinates $\tau^1 \in I^1 = [\tau_0, \tau_e]$ and $\tau^2 \in I^2 = [\tau_e, \tau_f]$ are related to $t \in [-1, 1]$ by the following linear transformations:

$$\tau^1 = \frac{(\tau_e - \tau_0)t + (\tau_e + \tau_0)}{2}$$
$$\tau^2 = \frac{(\tau_f - \tau_e)t + (\tau_f + \tau_e)}{2}$$

This results in the following reformulation of Eqs. (2.2-2.5)

$$\begin{aligned} J[\mathbf{x}(\cdot), \mathbf{u}(\cdot), \tau_0, \tau_e, \tau_f] = {} & E(\mathbf{x}_0, \mathbf{x}_e^-, \mathbf{x}_e^+, \mathbf{x}_f; \tau_0, \tau_e, \tau_f) \\ & + \frac{\tau_e - \tau_0}{2} \int_{-1}^{1} F^1(\mathbf{x}(t), \mathbf{u}(t), \tau(t)) dt \\ & + \frac{\tau_f - \tau_e}{2} \int_{-1}^{1} F^2(\mathbf{x}(t), \mathbf{u}(t), \tau(t)) dt \end{aligned} \tag{3.1}$$

$$\begin{aligned} \mathbf{f}_l^1 \le \mathbf{f}^1((\frac{2}{\tau_e - \tau_0})\dot{\mathbf{x}}(t), \mathbf{x}(t), \mathbf{u}(t), \tau(t)) \le \mathbf{f}_u^1 \\ \mathbf{f}_l^2 \le \mathbf{f}^2((\frac{2}{\tau_f - \tau_e})\dot{\mathbf{x}}(t), \mathbf{x}(t), \mathbf{u}(t), \tau(t)) \le \mathbf{f}_u^2 \end{aligned} \tag{3.2}$$

$$\mathbf{e}_l \le \mathbf{e}\big(\mathbf{x}_0, \mathbf{x}(\tau_e^-), \mathbf{x}(\tau_e^+), \mathbf{x}_f, \tau_0, \tau_e, \tau_f\big) \le \mathbf{e}_u \tag{3.3}$$

$$\begin{aligned} \mathbf{g}_l^1 \le \mathbf{g}^1(\mathbf{u}(t), \mathbf{x}(t), \tau(t)) \le \mathbf{g}_u^1 \\ \mathbf{g}_l^2 \le \mathbf{g}^2(\mathbf{u}(t), \mathbf{x}(t), \tau(t)) \le \mathbf{g}_u^2 \end{aligned} \tag{3.4}$$

Strictly speaking, we must use different symbols for all the mappings due to the transformation of the domain. However, we abuse notation here and retain these symbols for the purpose of brevity. Thus, in this context, one must view $\mathbf{x}(t)$ in the first segment, for example, as

$$\mathbf{x}(\tau(t)) = \mathbf{x}\left(\frac{(\tau_e - \tau_0)t + (\tau_e + \tau_0)}{2}\right)$$

for $\tau \in I^1$.

3.1 *Problem Discretization*

In the Legendre pseudospectral method, the LGL node points are closely related to the Legendre polynomials which are orthogonal over the interval $[-1, 1]$ with respect to a unit weight function. Let $L_N(t)$ be the Legendre polynomial of degree N on the interval $[-1, 1]$. The LGL points $t_l, l = 0, \ldots, N$ are given by

$$t_0 = -1, \quad t_N = 1$$

and for $1 \le l \le N-1$, t_l are the zeros of $\dot{L}_N$, the derivative of the Legendre polynomial, L_N. The method may now be elaborated as follows: Let the integers $N^1 + 1$ and $N^2 + 1$ denote the number of LGL points on these subintervals. The approximate solutions on these intervals I^1 and I^2 are denoted by $\mathbf{x}_N^i$ for states and $\mathbf{u}_N^i$ for controls, $i = 1, 2$

$$\mathbf{x}(\tau) \approx (\mathbf{x}_N^1(\tau), \mathbf{x}_N^2(\tau))$$
$$\mathbf{u}(\tau) \approx (\mathbf{u}_N^1(\tau), \mathbf{u}_N^2(\tau))$$

where the subscript N is used instead of N^i when it is clear form the context whether it is N^1 or N^2. The approximate states and controls are assumed to be a linear combination of Lagrange interpolating polynomials, $\phi_l(t)$

$$\mathbf{x}_N^1(\tau^1) = \sum_{l=0}^{N^1} \mathbf{x}^1(\tau_l^1)\phi_l(t), \mathbf{x}_N^2(\tau^2) = \sum_{l=0}^{N^2} \mathbf{x}^2(\tau_l^2)\phi_l(t)$$
$$\mathbf{u}_N^1(\tau^1) = \sum_{l=0}^{N^1} \mathbf{u}^1(\tau_l^1)\phi_l(t), \mathbf{u}_N^2(\tau^2) = \sum_{l=0}^{N^2} \mathbf{u}^2(\tau_l^2)\phi_l(t)$$

where t is in the computational domain $[-1, 1]$ and t_l for $l = 0, \ldots, N^i, i = 1, 2$ are the LGL points. The node points in I^1 and I^2 are denoted by τ_l^1 and τ_l^2, respectively. The Lagrange interpolating polynomials which interpolate the functions at the LGL points are for $l = 0, 1, ..., N^i$, $i = 1, 2$

$$\phi_l(t) = \frac{1}{N^i(N^i+1)L_{N^i}(t_l)} \frac{(t^2-1)\dot{L}_{N^i}(t)}{t - t_l} \quad (3.5)$$

To carryout the discretization of the problem, we impose the condition that the approximations above satisfy the differential inclusions at the LGL node points. To express the derivative $\dot{\mathbf{x}}_N^i(t)$ in terms of $\mathbf{x}_N^i(t)$ at the node points t_k, we differentiate the approximate solutions and evaluate the result at t_k to obtain a matrix multiplication of the following form for $i = 1, 2$:

$$\dot{\mathbf{x}}_N^i(t_k) = \sum_{l=0}^{N^i} \mathbf{x}^i(t_l)\dot{\phi}_l(t_k) = \sum_{l=0}^{N^i} D_{kl}\mathbf{x}^i(t_l) \quad (3.6)$$

where $D_{kl} = \dot{\phi}_l(t_k)$ are entries of the $(N^i+1) \times (N^i+1)$ differentiation matrix $\mathbf{D}$

$$\mathbf{D} := [D_{kl}] := \begin{cases} \dfrac{L_{N^i}(t_k)}{L_{N^i}(t_l)} . \dfrac{1}{t_k - t_l} & k \neq l \\ -\dfrac{N^i(N^i+1)}{4} & k = l = 0 \\ \dfrac{N^i(N^i+1)}{4} & k = l = N^i \\ 0 & \text{otherwise} \end{cases} \quad (3.7)$$

In terms of the following coefficients

$$\begin{aligned} \mathbf{X} \equiv (\mathbf{X}^1, \mathbf{X}^2) = (&\mathbf{x}^1(\tau_0^1), \mathbf{x}^1(\tau_1^1), \ldots, \mathbf{x}^1(\tau_{N^1}^1), \\ &\mathbf{x}^2(\tau_0^2), \mathbf{x}^2(\tau_1^2), \ldots, \mathbf{x}^2(\tau_{N^2}^2)) \\ \mathbf{U} \equiv (\mathbf{U}^1, \mathbf{U}^2) = (&\mathbf{u}^1(\tau_0), \mathbf{u}^1(\tau_1^1), \ldots, \mathbf{u}^1(\tau_{N^1}^1), \\ &\mathbf{u}^2(\tau_0^2), \mathbf{u}^2(\tau_1^2), \ldots, \mathbf{u}^2(\tau_{N^2}^2)) \end{aligned}$$

the discretization of the dynamic constraints can be carried out in the following way:

$$\begin{aligned} \mathbf{f}_l^1 \leq \mathbf{f}^1((\frac{2}{\tau_e - \tau_0})\dot{\mathbf{x}}_{1,k}, \mathbf{x}_{1,k}, \mathbf{u}_{1,k}, \tau_k^1) \leq \mathbf{f}_u^1 \\ \text{for } k = 0, \ldots, N^1 \\ \mathbf{f}_l^2 \leq \mathbf{f}^2((\frac{2}{\tau_f - \tau_e})\dot{\mathbf{x}}_{2,k}, \mathbf{x}_{2,k}, \mathbf{u}_{2,k}, \tau_k^2) \leq \mathbf{f}_u^2 \\ \text{for } k = 0, \ldots, N^2 \end{aligned} \quad (3.8)$$

where the simplified notation $\mathbf{x}_{i,k} = \mathbf{x}^i(\tau_k^i), \mathbf{u}_{i,k} = \mathbf{u}^i(\tau_k^i)$ is used.

Next, using the Gauss-Lobatto integration rule (quadrature) the Bolza cost function in (3.1) is discretized and the integrals are approximated by a finite sum.

$$\begin{aligned} J^N(\mathbf{X}, \mathbf{U}, \tau_0, \tau_f) = E(&\mathbf{x}_{1,0}, \mathbf{x}_{1,N^1}, \mathbf{x}_{2,0}, \mathbf{x}_{2,N_2}; \tau_0, \tau_e, \tau_f) \\ &+ \frac{\tau_e - \tau_0}{2} \sum_{k=0}^{N^1} F^1(\mathbf{x}_{1,k}, \mathbf{u}_{1,k}, t_k) w_k^1 \\ &+ \frac{\tau_f - \tau_e}{2} \sum_{k=0}^{N^2} F^2(\mathbf{x}_{2,k}, \mathbf{u}_{2,k}, t_k) w_k^2 \end{aligned} \quad (3.9)$$

In the above w_k^i are the weights given by

$$w_k^i := \frac{2}{N^i(N^i+1)} \frac{1}{[L_{N^i}(t_k)]^2} \ i = 1, 2, \ k = 0, 1, \ldots, N \quad (3.10)$$

and $\mathbf{x}_{1,N^1} = \mathbf{x}_e^-, \mathbf{x}_{2,0} = \mathbf{x}_e^+$. The event and the path constraints are also discretized as follows:

$$\mathbf{e}_l \leq \mathbf{e}\big(\mathbf{x}_{1,0}, \mathbf{x}_{1,N^1}, \mathbf{x}_{2,0}, \mathbf{x}_{2,N^2}; \tau_0, \tau_e, \tau_f\big) \leq \mathbf{e}_u \quad (3.11)$$

$$\mathbf{g}_l^1 \leq \mathbf{g}^1(\mathbf{x}_{1,k}, \mathbf{u}_{1,k}, \tau_k^1) \leq \mathbf{g}_u^1, \ k = 0, \ldots, N^1 \quad (3.12)$$
$$\mathbf{g}_l^2 \leq \mathbf{g}^2(\mathbf{x}_{2,k}, \mathbf{u}_{2,k}, \tau_k^2) \leq \mathbf{g}_u^2, \ k = 0, \ldots, N^2 \quad (3.13)$$

The optimal control problem is thus approximated to the NLP of finding coefficients $\mathbf{X} = (\mathbf{X}^1, \mathbf{X}^2), \mathbf{U} = (\mathbf{U}^1, \mathbf{U}^2)$ and possibly the event times τ_0, τ_e, τ_f that minimize the cost function (3.9) subject to the constraints Eq. (3.8) and Eqs. (3.11)-(3.13).

3.2 *Knots and Nodes*

At this point, it is necessary to clarify the terminology of *hard* and *soft* knots which are integral to the above formulation and also delineate the idea of the *event constraints* more clearly. First note that $\tau_{N^1}^1 = \tau_0^2 = \tau_e$. We use the word *knots* to denote such points. The "knotting condition" given by Eq. (3.11) transfers information at the knots across the two segments 1 and 2. We call such nodes hard knots and they are *intrinsic* to the problem formulation. For example, the dropping of a stage (i.e.

mass discontinuity) in a two-stage launch problem defines a hard knot. Suppose now that the problem did not have the interior event. We can hence define a "false event" of state continuity at an arbitrary point τ_e,

$$\mathbf{x}_e^- \equiv \mathbf{x}_e^+$$

leading to a linear knotting condition

$$\mathbf{x}_{1,N^1} - \mathbf{x}_{2,0} = \mathbf{0} \tag{3.14}$$

We call such nodes, *soft knots*. These knots can be used to enhance pseudospectral methods in various ways. In the present context, they are used to capture nonsmooth behavior like switches in the control. We use the terms *fixed* (hard/soft) knot if τ_e is fixed and *free* (hard/soft) knot if τ_e is free. Clearly, fixed/free hard knots are based on problem formulation and fixed/free soft knots are "designer knots".

The pseudospectral knotting method is implemented in a reusable software package called DIDO (Ross and Fahroo, 2001). In the following section, we discuss the application of soft knots for capturing nonsmoothness in an example problem.

4. A NUMERICAL EXAMPLE

The Moon-Landing problem is a very simple, yet illustrative problem (Meditch, 1964). The problem is to minimize fuel during the vertical descent of a lander. It can be formulated as "smooth" optimal control problem. However, a hodograph transformation allows the elimination of the control variable while reposing the dynamic constraints as a differential inclusion (Clarke, 1990). We use the differential inclusion formulation to illustrate the ability of our method in solving such problems (Fahroo and Ross, 2001a).

The problem is a *time-free* problem of maximizing the final mass or minimizing

$$J = -m(\tau_f) \tag{4.1}$$

subject to the dynamic constraints

$$\frac{dh}{d\tau} - v = 0, \tag{4.2}$$

$$0 \le m\frac{dv}{d\tau} + mg \le T_{max}, \tag{4.3}$$

$$-\frac{T_{max}}{I_{sp}g} \le \frac{dm}{d\tau} \le 0 \tag{4.4}$$

where the state variables h, v and m are altitude, speed and mass, respectively. The constant parameters in the problem are g, the gravity of moon (or any planet without an atmosphere), I_{sp}, the specific impulse of the propellent and $T_{max} > 0$ the maximum value of the thrust, T. The hard knots for this problem are given by the boundary conditions. The initial conditions were *arbitrarily chosen* to be

$$h(0) = 1.0, \quad v(0) = -0.05, \quad m(0) = 1.0. \tag{4.5}$$

while the final conditions are given by the soft landing requirement

$$h(\tau_f) = 0, \quad v(\tau_f) = 0. \tag{4.6}$$

and a physically realizable trajectory,

$$m(\tau_f) > 0 \tag{4.7}$$

The last equation is numerically imposed as $m(\tau_f) \ge \epsilon$ where ϵ is the machine precision. Thus, Eqs.(4.5)-(4.7) are the *event constraints*. The constant parameters for this problem were arbitrarily chosen as

$$\frac{T_{max}}{m_0 g} = 0.5, \quad \frac{I_{sp}g}{v_0} = 1$$

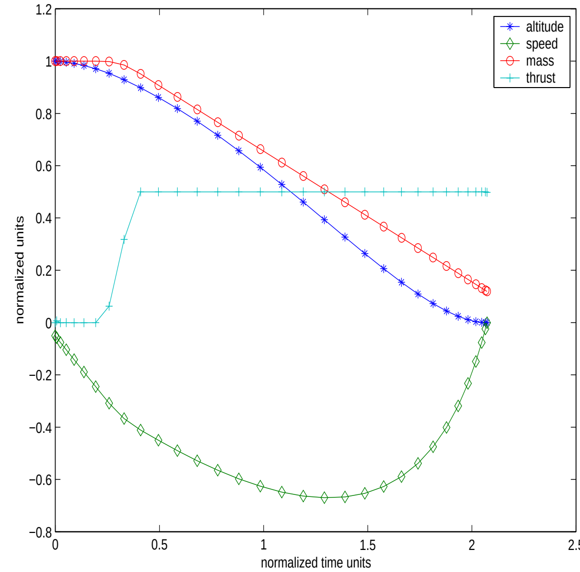

Figure 1: State and Control Histories With Smooth Method

The graphs of state and control variables obtained by implementing the smooth pseudospectral method for 32 LGL points is shown in Figure 1. The control, $T = -\dot{m}I_{sp}g$, rapidly changes around $t = 0.5$ suggesting the possibility of a switch. When one free soft knot is introduced, the nonsmooth characteristics are readily apparent as seen in Figure 2. The number of nodes is still 32; however, they are now split to 8 nodes on the first interval while the second interval contains 24 points. It is apparent that the smooth method smears the discontinuity in the thrust – an artifact of the Gibbs phenomenon, while the nonsmooth (knotting) method works well in capturing the nondifferentiabilities in the state and control variables.

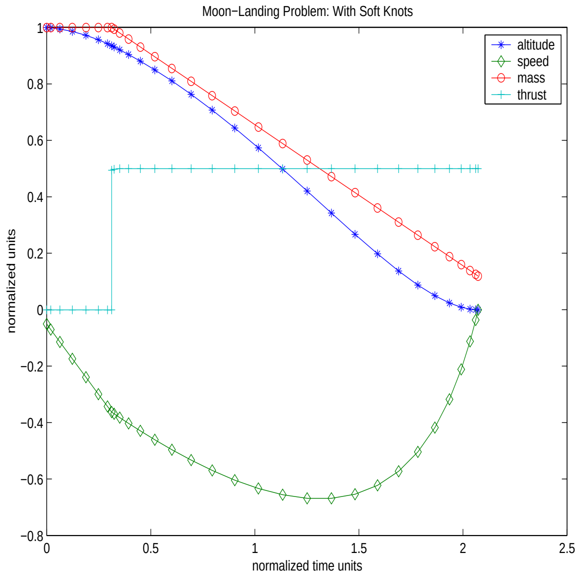

Figure 2: State and Control Histories With Soft Knots

It is worth noting that m and v are nondifferentiable at precisely the switch point which is the soft knot. Further, since $-\dot{m} = T/(I_{sp}g)$, it is apparent that by regarding $0 \leq T \leq T_{max}$, we must interpret $-\dot{m}$ at the switch point as the generalized derivative, $-\partial_t m$ (Vinter, 1990).

5. CONCLUSIONS

Knotting methods based on a Legendre pseudospectral method for direct trajectory optimization is proposed. Its extension to other spectral methods, such as those based on Chebyshev polynomials, is quite straightforward although the Legendre method offers the most natural choice for the discretization of a generalized Bolza optimal control problem. This method offers great flexibility in solving nonsmooth optimal control problems. In addition, our proposed method can efficiently handle a vast number of complexities arising in real-world optimal control problems such as rapid changes in dynamics, state-dependent control constraints and event conditions. A nonlinear model-predictive controller appears to be within reach by way of real-time optimization. However, further theoretical analysis and numerical experimentations are necessary to fully exploit this approach.

6. ACKNOWLEDGMENTS

We gratefully acknowledge the financial support provided in part by Draper Labs, Jet Propulsion Laboratory and the Naval Postgraduate School.

7. REFERENCES

Betts, J. T., (1998). Survey of numerical methods for trajectory optimization. *Journal of Guidance, Control, and Dynamics,* Vol. 21, No. 2, 193–207.

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

Clarke, F. H. (1990). *Optimization and Nonsmooth Analysis*, Classics in Applied Mathematics, Vol. 5, SIAM, Philadelphia, PA.

Elnagar, J., M. A. Kazemi and M. Razzaghi (1995). The pseudospectral Legendre method for discretizing optimal control problems. *IEEE Transactions on Automatic Control,* Vol. 40, No. 10, 1793–1796.

Gottlieb, D., M. Y. Hussaini and S. A. Orszag (1984). Theory and applications of spectral methods. *Spectral Methods for PDE's,* SIAM, Philadelphia, PA.

Fahroo, F. and I. M. Ross (2001a). Second look at approximating differential inclusions. *Journal of Guidance, Control and Dynamics*, Vol. 24, No. 1, 131–133.

Fahroo, F. and I. M. Ross (2001b). Costate estimation by a Legendre pseudospectral method. *Journal of Guidance, Control, and Dynamics*, Vol. 24, No. 2, 270–277.

Fornberg, B. (1998). *A Practical Guide to Pseudospectral Methods.* Cambridge University Press, New York, N.Y.

Meditch, J. S. (1964). On the problem of optimal thrust programming for a lunar soft landing. *IEEE Transactions on Automatic Control*, Vol. AC-9, 477-484.

Ross, I. M. and F. Fahroo (2001). *User's Manual for DIDO 2001: A MATLAB Application Package for Dynamic Optimization.* NPS Technical Report AA-01-003, Department of Aeronautics and Astronautics, Naval Postgraduate School, Monterey, CA.

Vinter, R. (2000). *Optimal Control.* Birkhäuser, Boston, MA.