Optimization of Higher-Order Systems and Extensions of Minimum Principle

Tawiwat Veeraklaew

Department of Mechanical Engineering, Chulachomklao Royal Military Academy
Maung, Nakhon-Nayok, 26001
Tel: (037) 393010-4 ext 62305, Fax: (037) 393487, E-mail: tawiwat@hotmail.com

Abstract

In recent years, using tools of linear and nonlinear systems theory, it has been shown that a large number of dynamic systems can be written in canonical forms. These canonical forms are alternatives to state-space forms and can be represented by higher-order differential equations. For planning and control purposes, these canonical forms provide a number of advantages when compared to their corresponding first-order forms. In this paper, we address the question of optimization of dynamic systems described by higher-order differential equations. The minimum principle for higher-order systems is derived directly from their higher-order form and the results are confirmed by classical theory using first-order form.

In this paper, optimality conditions for higher-order systems are derived using two approaches: (i) Hamilton-Jacobi theory, thereby, extending Pontryagin's principle; (ii) variational calculus of higher-order augmented cost functional. It is shown that the two approaches lead to the same results. Further, the results of these two approaches are also verified by variational calculus of their equivalent first-order forms. The result applicable to higher-order systems is illustrated by an example.

1 Introduction

$-L_c\hat{z}_i$ It is customary to express dynamic systems in state-space form, i.e., a system of first-order differential equations. However, for a large class of mechanical systems, the most natural representation of the system is a set of second-order differential equations which arise from the application of Newton's laws. Needless to say, these second-order differential equations can be converted to a set of first-order differential equations but this process is accompanied by inversion of the inertia matrix that makes the first-order form unreasonably complicated ([21], [5]). From a different perspective, using the theory of linear and nonlinear systems, see e.g. [13], tools of differential geometry provide alternate representations of systems in canonical forms which allow the system to be rewritten as higher-order differential equations in the canonical variables. Dynamic systems that have this feature include controllable linear systems [1], feedback linearizable systems [13], chained form systems [16], and differentially flat systems [10].

This paper addresses the underlying theory for optimal trajectory generation for this broad class of systems which have a higher-order representation in their original coordinates or in the transformed coordinates. Currently, methods which exploit the structure of the higher-order differential equations to efficiently compute the optimal solution are still in their infancy. In a recent study, a direct method was used for a class of higher-order systems to compute the optimal solution [17]. In some recent works, Agrawal and coworkers have exploited the structure of the higher-order equations to compute the optimal solution using indirect methods. These studies assumed no inequality constraints and the state-space equations were explicitly embedded into the cost functional, thereby, reducing a constrained optimization problem to an unconstrained optimization problem. The results from higher-order variational calculus were used to find the necessary conditions for optimality [12]. This approach was demonstrated for linear time-invariant systems [1], classes of time-varying systems [2], feedback linearizable systems [3], and fully actuated robot systems [4]. A recent study uses the explicit structure of globally feedback linearizable systems to derive some important results applicable to optimal solution of Mayer's problem in the presence of inequality constraints [20]. At this time, the authors are unaware of a unifying theory that would extend the minimum principle [18] to systems in higher-order form, without converting them

to first-order form. Equivalence between the results of the minimum principle and calculus of variations with the first-order form have been demonstrated in the literature ([7], [6], [19], [15], [9]).

The purpose of this paper is to extend the classical results of the minimum principle, which apply only to systems described in first-order forms, to systems in higher-order forms. For completeness, these results are derived using Hamilton-Jacobi theory as well as calculus of variations applicable to functionals with high derivatives. Further, these results are verified using the classical results of minimum principle by rewriting the higher-order system in a first-order form. The results obtained by these alternative approaches are then compared to gain a better insight into their equivalence.

The statement of the problem is to find the optimal trajectory for a dynamic system described by

$$x^{(p)}(t) = f(x, x^{(1)}, ..., x^{(p-1)}, u), \tag{1}$$

where $x \in \mathcal{R}^n$ and $u \in \mathcal{R}^m$ and $x^{(i)}$ represents the ith derivative of x. The trajectory must minimize the cost

$$J = \Phi(x(t_f), x^{(1)}(t_f), ..., x^{(p-1)}(t_f), t_f) + \int_{t_0}^{t_f} L(x(\tau), x^{(1)}(\tau), ..., x^{(p-1)}(\tau), u(\tau), \tau) d\tau \tag{2}$$

and satisfy the following constraints: (i) $u(t) \in \mathcal{A}[t, x(t), x^{(1)}(t), ..., x^{(p-1)}(t)]$; (ii) Initial conditions of the trajectories $x(t)$, $x^{(1)}(t)$,...,$x^{(p-1)}(t)$ are specified while the terminal conditions are free.

Some special cases of this problem are: (i) Here, the set $u \in \mathcal{A}[t, x(t), x^{(1)}(t), ..., x^{(p-1)}(t)]$ represents general equality and inequality constraints on the actuator inputs. A special case of this constraint is when $\mathcal{A}[t, x(t), x^{(1)}(t), ..., x^{(p-1)}(t)]$ is independent of states and time; (ii) The cost functional (2) includes some important problems such as minimum time, minimum fuel, maximum terminal states as special cases; (iii) If $p = 1$, Eq. (1) is the familiar state space description of a system. Optimization of this class has been dealt with extensively in the literature.

The purpose of this paper is to extend the theory of dynamic optimization to the case $p > 1$. In this paper, the optimization theory will be derived using two approaches: (i) Hamilton-Jacobi theory to arrive at the extended form of Pontryagin's principle that applies to higher-order systems; (ii) variational calculus after the inequality constraints of the set $\mathcal{A}[t, x(t), x^{(1)}(t), ..., x^{(p-1)}(t)]$ have been transformed to equality constraints. It will be demonstrated that the two approaches lead to the same results. Also, these results are consistent with those obtained by the use of classical minimum principle applied to first-order form. The organization of this paper is as follows: Section 2 describes the Hamilton-Jacobi approach to derive the optimality equations while Section 3 addresses the same problem using variational theory when the system is in the higher-order form or is expressed in the first-order form. The results of this theory are illustrated with an example of a nonliner system in Section 4.

2 Hamilton-Jacobi Theory

In this section, the derivation of the extended form of Pontryagin's principle is performed in two steps. First, it is considered that the constraint set is independent of time and states, i.e., $\mathcal{A}[t, x(t), x^{(1)}(t), ..., x^{(p-1)}(t)] = \mathcal{U}$. The optimality conditions are derived for this case using the pattern of proof given in [14]. Second, the constraint set is considered to be time and state dependent, i.e., $\mathcal{A}[t, x(t), x^{(1)}(t), ..., x^{(p-1)}(t)]$. The optimality conditions are extended for this case. In order to abbreviate notations, we define $\mathbf{x}(t)^T = (x(t)^T\ x^{(1)}(t)^T\ ...\ x^{(p-1)}(t)^T)^T$.

2.1 Constraint Set $\mathcal{U}$

We imbed this problem in a larger problem class by defining a return function

$$J(\mathbf{x}(t), t, u(\tau)) = \Phi(\mathbf{x}(t_f), t_f) + \int_t^{t_f} L(\mathbf{x}(\tau), u(\tau), \tau) d\tau \tag{3}$$

Here, $\mathbf{x}(t)$ is an admissible start point at t and $u(\tau)$ is an admissible input trajectory, i.e., $u(\tau) \in \mathcal{U}$ and is defined over $t \leq \tau \leq t_f$. For this start point $(t, \mathbf{x}(t))$, we define the minimum cost

$$J^*(\mathbf{x}(t), t) = \min_{u(\tau) \in \mathcal{U}} \{\Phi(\mathbf{x}(t_f), t_f) + \int_t^{t_f} L(\mathbf{x}(\tau), u(\tau), \tau) d\tau\}. \tag{4}$$

By subdividing the interval (t, t_f) into $t \leq \tau \leq t+\Delta t$ and $t+\Delta t < \tau \leq t_f$, from the principle of optimality, we can rewrite Eq. (4) as

$$J^*(\mathbf{x}(t), t) = \min_{u(\tau) \in \mathcal{U}} \{\int_t^{t+\Delta t} L d\tau + J^*(\mathbf{x}(t+\Delta t), t+\Delta t)\} \tag{5}$$

On expanding Eq. (5) in Taylor's series about $(\mathbf{x}(t), t)$, we obtain

$$J^*(\mathbf{x}(t), t) = \min_{u(\tau) \in \mathcal{U}} \{\int_t^{t+\Delta t} L d\tau + J^*(\mathbf{x}(t), t)$$

$$+\frac{\partial J^*}{\partial t}(\mathbf{x}(t),t)\Delta t+[\frac{\partial J^*}{\partial \mathbf{x}}(\mathbf{x}(t),t)]^T\dot{\mathbf{x}}(t)\Delta t +h.o.t.\} \quad (6)$$

Recalling $\mathbf{x}(t)^T = (x(t)^T\ x^{(1)}(t)^T\ ...\ x^{(p-1)}(t)^T)^T$ and for small Δt,

$$\begin{aligned}J^*(\mathbf{x}(t),t) &= \min_{u(t)\in\mathcal{U}}\{J^*(\mathbf{x}(t),t)\\ &+[L(\mathbf{x}(t),u(t),t)+J_t^*(\mathbf{x}(t),t)+\\ &J_x^{*T}(\mathbf{x}(t),t)x^{(1)}(t)+...\\ &+J_{x^{(p-2)}}^{*T}(\mathbf{x}(t),t)x^{(p-1)}(t)+\\ &J_{x^{(p-1)}}^{*T}(\mathbf{x}(t),t)f(\mathbf{x}(t),u(t),t)]\Delta t+o(\Delta t)\} \quad (7)\end{aligned}$$

On neglecting higher-order terms of Δt and separating terms dependent on $u(t)$, we can rewrite the above equation as

$$\begin{aligned}0 &= J_t^*(\mathbf{x}(t),t)+J_x^{*T}(\mathbf{x}(t),t)x^{(1)}(t)+...\\ &+J_{x^{(p-2)}}^{*T}(\mathbf{x}(t),t)x^{(p-1)}(t)\\ &+\min_{u(t)\in\mathcal{U}}\{L(\mathbf{x}(t),u(t),t)\\ &+J_{x^{(p-1)}}^{*T}(\mathbf{x}(t),t)f(\mathbf{x}(t),u(t),t)\}\end{aligned} \quad (8)$$

One can now define a Hamiltonian $\mathcal{H}$

$$\begin{aligned}\mathcal{H}(\mathbf{x}(t),u(t),J_{x^{(p-1)}}^*,t) &= L(\mathbf{x}(t),u(t),t)\\ &+J_{x^{(p-1)}}^{*T}(\mathbf{x}(t),t)f(\mathbf{x}(t),u(t),t)\end{aligned} \quad (9)$$

and

$$\begin{aligned}\mathcal{H}(\mathbf{x}(t),u^*(\mathbf{x}(t),J_{x^{(p-1)}}^*,t),J_{x^{(p-1)}}^*,t) &=\\ \min_{u(t)\in\mathcal{U}}\mathcal{H}(\mathbf{x}(t),u(t),J_{x^{(p-1)}}^*,t)\end{aligned} \quad (10)$$

Here, the minimizing control is said to depend on $\mathbf{x}(t)$, $J_{x^{(p-1)}}^*(\mathbf{x}(t),t)$, and t. From Eqs. (8) and (10), the extended form of Hamilton-Jacobi equation is

$$\begin{aligned}&J_t^*(\mathbf{x}(t),t)+J_x^{*T}(\mathbf{x}(t),t)x^{(1)}(t)+...\\ &+J_{x^{(p-2)}}^{*T}(\mathbf{x}(t),t)x^{(p-1)}(t)\\ &+\mathcal{H}(\mathbf{x}(t),u^*(\mathbf{x}(t),J_{x^{(p-1)}}^*,t),J_{x^{(p-1)}}^*,t)\\ &=0,\end{aligned} \quad (11)$$

where J satisfies the boundary condition

$$J^*(\mathbf{x}(t_f),t_f)=\Phi(\mathbf{x}(t_f),t_f). \quad (12)$$

In summary, (i) the optimal control $u^*(\mathbf{x}(t),J_{x^{(p-1)}}^*,t)$ minimizes the Hamiltonian $\mathcal{H}$ defined in Eq. (10), and (ii) the optimal return function satisfies the partial differential equation (11). If $p=1$, these results simplify to the the classical Hamilton-Jacobi equations. Eq. (10) is an equivalent statement made by Pontryagin [18].

2.1.1 Extended Pontryagin's Principle

The extended Pontryagin's principle for the case of $p>1$ can be derived from Hamilton-Jacobi Eq. (11) using the pattern suggested by Kirk [14]. If $(\mathbf{x}^*(t),t)$ is a point on the optimal trajectory, Hamilton-Jacobi equation can also be written as

$$\begin{aligned}0 &= \min_{u(t)\in\mathcal{U}}\{J_t^*(\mathbf{x}^*(t),t)+J_x^{*T}(\mathbf{x}^*(t),t)x^{*(1)}(t)\\ &+...+J_{x^{(p-2)}}^{*T}(\mathbf{x}^*(t),t)x^{*(p-1)}(t)+\\ &\mathcal{H}(\mathbf{x}^*(t),u(t),J_{x^{(p-1)}}^*,t)\},\end{aligned} \quad (13)$$

since $J_t^*(\mathbf{x}^*(t),t)$, $J_x^{*T}(\mathbf{x}^*(t),t)x^{*(1)}(t)$, ..., $J_{x^{(p-2)}}^{*T}(\mathbf{x}^*(t),t)x^{*(p-1)}(t)$ are independent of $u(t)$. In words, for a $(\mathbf{x}^*(t),t)$, the control $u^*(t)$ minimizes the right hand side of Eq. (13) and the minimum is zero. Hence, if we define a function

$$\begin{aligned}v(\mathbf{x}(t),u^*(t),t) &= J_t^*(\mathbf{x}(t),t)+\\ &J_x^{*T}(\mathbf{x}(t),t)x^{(1)}(t)+...\\ &+J_{x^{(p-2)}}^{*T}(\mathbf{x}(t),t)x^{(p-1)}(t)+\\ &\mathcal{H}(\mathbf{x}(t),u^*(t),J_{x^{(p-1)}}^*,t),\end{aligned} \quad (14)$$

in the neighborhood of $\mathbf{x}^*(t)$, i.e., $\mathbf{x}(t)=\mathbf{x}^*(t)+\delta\mathbf{x}(t)$, this function has a local mimimum at $\mathbf{x}^*(t)$. If $\mathbf{x}(t)$ is not constrained by any boundaries, this local minimum property can be written mathematically as $\frac{\partial v}{\partial \mathbf{x}}(\mathbf{x}^*(t),u^*(t),t)=0$. Assuming that the mixed partial derivatives are continuous, the order of the derivatives in a mixed partial can be interchanged. With this property, $\frac{\partial v}{\partial \mathbf{x}}(\mathbf{x}^*(t),u^*(t),t)=0$ simplifies to the following component equations:

$$\begin{aligned}&J_{x^{(k)}t}^*+J_{x^{(k)}x}^*x^{(1)}+...+J_{x^{(k)}x^{(k-1)}}^*x^{(k)}\\ &+J_{x^{(k-1)}}^*+...+J_{x^{(k)}x^{(p-2)}}^*x^{(p-1)}\\ &+J_{x^{(k)}x^{(p-1)}}^*x^{(p)}+L_{x^{(k)}}+f_{x^{(k)}}^TJ_{x^{(p-1)}}^*\\ &=0,\qquad k=0,...,p-1,\end{aligned} \quad (15)$$

evaluated at $\mathbf{x}^*(t)$, $u^*(t)$, t. Here, a term such as $J_{x^{(k)}x^{(p-2)}}^*$ denotes a $(n\times n)$ matrix with rs element $J_{x_r^{(k)}x_s^{(p-2)}}^*$. Using the definition of total time derivative of function $J_{x^{(k)}}^*$, the above equation simplifies to

$$\begin{aligned}&\frac{\mathrm{d}J_{x^{(k)}}^*}{\mathrm{d}t}+J_{x^{(k-1)}}^*+L_{x^{(k)}}+f_{x^{(k)}}^TJ_{x^{(p-1)}}^*=0,\\ &k=0,...,p-1.\end{aligned} \quad (16)$$

On defining $\psi_k(t)=J_{x^{(k-1)}}^*(\mathbf{x}^*(t),t)$, Eq. (16) can be written as

$$\begin{aligned}&\psi_{k+1}^{(1)}(t)+\psi_k(t)+\mathcal{H}_{x^{(k)}}(\mathbf{x}^*(t),u^*(t),t)=0,\\ &k=0,...,p-1\end{aligned} \quad (17)$$

Note that ψ_k are defined for values 1 through p. In summary, each $(\mathbf{x}^*(t), t)$ on the optimal path satisfies Eq. (17). The components of this equation are

$$\begin{aligned}
&\psi_p^{(1)}(t) + \psi_{p-1}(t) + \mathcal{H}_{x^{(p-1)}}(\mathbf{x}^*(t), u^*(t), t) = 0 \\
&\psi_{p-1}^{(1)}(t) + \psi_{p-2}(t) + \mathcal{H}_{x^{(p-2)}}(\mathbf{x}^*(t), u^*(t), t) = 0 \\
&\vdots \\
&\psi_2^{(1)}(t) + \psi_1(t) + \mathcal{H}_{x^{(1)}}(\mathbf{x}^*(t), u^*(t), t) = 0 \\
&\psi_1^{(1)}(t) + \mathcal{H}_x(\mathbf{x}^*(t), u^*(t), t) = 0 \qquad (18)
\end{aligned}$$

Using this equations, it is possible to eliminate $\psi_{p-1}(t)$ through $\psi_1(t)$. Then, the resulting differential equation is

$$\mathcal{H}_x - \mathcal{H}_{x^{(1)}}^{(1)} + ... + (-1)^{p-1}\mathcal{H}_{x^{(p-1)}}^{(p-1)} = (-1)^p \psi_p^{(p)} \quad (19)$$

which must hold at each point $(\mathbf{x}^*(t), t)$ of the optimal solution.

The solution of this differential equation requires boundary conditions $\psi_p(t_f)$,...,$\psi_p^{(p-1)}(t_f)$. Using Eq. (12) and from the definition,

$$\psi_k(\mathbf{x}^*(t_f), t_f) = \Phi_{x^{(k-1)}}(\mathbf{x}^*(t_f), t_f), \quad k = 1, ..., p \qquad (20)$$

From Eqs. (18), it can be shown that

$$\begin{aligned}
&\psi_{p-k} = (-1)^k \psi_p^{(k)} + (-1)^k \mathcal{H}_{x^{(p-1)}}^{(k-1)} \\
&+(-1)^{k-1}\mathcal{H}_{x^{(p-2)}}^{(k-2)} + ... + (-1)\mathcal{H}_{x^{(p-k)}}, \\
&\quad k = 1, ..., p-1 \qquad (21)
\end{aligned}$$

Given $\psi_{p-k}(t_f)$ from Eq. (20) and the expression of Hamiltonian, Eq. (21) can be used to compute the boundary values $\psi_p^{(k)}(t_f)$.

In summary, for a dynamic system described by differential Eqs. (1) with boundary conditions $x(0)$, $x^{(1)}(0)$,...,$x^{(p-1)}(0)$, the optimal trajectories that minimize the cost functional (2) with $u(t) \in \mathcal{U}$ satisfy the following conditions: (i) define a Hamiltonian $\mathcal{H}(\mathbf{x}(t), u(t), t) = L(\mathbf{x}(t), u(t), t) + \psi_p^T(t) f(\mathbf{x}(t), u(t), t)$, (ii) find the minimum of the Hamiltonian $\mathcal{H}(\mathbf{x}(t), u^*(t), t)$ within the set $u(t) \in \mathcal{U}$, (iii) find $\psi_p(t)$ that satisfy the differential Eq. (19) along with boundary conditions of $\psi_p^{(k)}(t_f)$ computed from Eq. (21). When $p = 1$, these results simplify to those of classical minimum principle [9].

2.2 Constraint Set $\mathcal{A}(t, \mathbf{x})$

In this section, the constraint set for $u(t)$ is both time and state dependent, i.e., $u(t) \in \mathcal{A}(t, \mathbf{x})$. In the derivation of Hamilton-Jacobi equations, the steps are same as of Section 2.1 with Eq. (10) replaced in the following way:

$$\begin{aligned}
&\mathcal{H}(\mathbf{x}(t), u^*(\mathbf{x}(t), J^*_{x^{(p-1)}}, t), J^*_{x^{(p-1)}}, t) = \\
&\min_{u(t) \in \mathcal{A}(t,\mathbf{x})} \mathcal{H}(\mathbf{x}(t), u(t), J^*_{x^{(p-1)}}, t) \qquad (22)
\end{aligned}$$

In the derivation of the extended Pontryagin's principle, the steps remain essentially the same. $\mathbf{x}^*(t)$ now minimzes $v(\mathbf{x}(t), u^*(t), t)$ subject to the constraint set $u(t) \in \mathcal{A}(t, \mathbf{x})$. If the constraints are written mathematically as

$$C_j(t, \mathbf{x}, u) \leq 0, \quad j = 1, ..., r, \qquad (23)$$

$\mathbf{x}^*(t)$ minimizes $v'(\mathbf{x}(t), u^*(t), t)$, where

$$\begin{aligned}
&v'(\mathbf{x}(t), u^*(t), t) = v(\mathbf{x}(t), u^*(t), t) \\
&+ \sum_{j=1}^{r} \mu_j (C_j(t, \mathbf{x}, u^*) + \xi_j^2). \qquad (24)
\end{aligned}$$

This local minimum property can be written mathematically as $\frac{\partial v'}{\partial \mathbf{x}}(\mathbf{x}^*(t), u^*(t), t) = 0$. Using a line of arguments similar to Section 2.1 and defining a modified Hamiltonian $\mathcal{H}' = \mathcal{H} + \sum_{j=1}^{r} \mu_j C_j(t, \mathbf{x}, u^*)$, one can shown that

$$\begin{aligned}
&\psi_{k+1}^{(1)}(t) + \psi_k(t) + \mathcal{H}'_{x^{(k)}}(\mathbf{x}^*(t), u^*(t), t) = 0, \\
&k = 0, ..., p-1 \qquad (25)
\end{aligned}$$

and $\mu_j \xi_j = \mu_j C_j(t, \mathbf{x}^*, u^*) = 0$, $j = 1, ..., r$. This second condition says that if the jth constraint is active, $\mu_j \geq 0$, otherwise $\mu_j = 0$. The costate Eqs. (19) now get modified to

$$\mathcal{H}'_x - \mathcal{H}'^{(1)}_{x^{(1)}} + ... + (-1)^{p-1}\mathcal{H}'^{(p-1)}_{x^{(p-1)}} = (-1)^p \psi_p^{(p)}, \quad (26)$$

which must hold at each point $(\mathbf{x}^*(t), t)$ of the optimal solution. The solution of ψ_{p-k} gets modified to

$$\begin{aligned}
&\psi_{p-k} = (-1)^k \psi_p^{(k)} + (-1)^k \mathcal{H}'^{(k-1)}_{x^{(p-1)}} \\
&+(-1)^{k-1}\mathcal{H}'^{(k-2)}_{x^{(p-2)}} + ... + (-1)\mathcal{H}'_{x^{(p-k)}}, \\
&\quad k = 1, ..., p-1 \qquad (27)
\end{aligned}$$

In summary, the optimal trajectories for a higher-order system with constraints $u(t) \in \mathcal{A}(t, \mathbf{x})$ satisfies Eqs. (1), (22), (26) along with the condition $\mu_j \geq 0$ if jth constraint is active, otherwise $\mu_j = 0$.

3 Variational Theory

First, some general results from higher-order variational theory will be stated. Almost all books on variational calculus consider functionals dependent on at most the first derivatives ([6], [7], [8], [9]). These results will then be specialized to the problem defined in Section 1.

3.1 Key Results

For a functional $J = \Psi(x(t_f), x^{(1)}(t_f), ..., x^{(p-1)}(t_f), t_f) + \int_{t_0}^{t_f} F(x(t), x^{(1)}(t), ..., x^{(p-1)}(t), x^p(t), t)dt$ with open end states and end time, it can be shown from the principles of variational calculus ([12], [11]),

$$\begin{aligned}
\delta J = &\int_{t_0}^{t_f} h^T[F_x - F^{(1)}_{x^{(1)}} + ... + (-1)^p F^{(p)}_{x^{(p)}}]dt \\
&+[h^T(F_{x^{(1)}} - F^{(1)}_{x^{(2)}} + ... + (-1)^{p-1}F^{(p-1)}_{x^{(p)}})]^{t_f}_{t_0} \\
&+[h^{(1)T}(F_{x^{(2)}} - F^{(1)}_{x^{(3)}} + ... + (-1)^{p-2}F^{(p-2)}_{x^{(p)}})]^{t_f}_{t_0} \\
&+... + [h^{(p-1)T}F_{x^{(p)}}]^{t_f}_{t_0} \\
&+[F - x^{(1)T}\{F_{x^{(1)}} + ... + (-1)^{p-1}F^{(p-1)}_{x^{(p)}}\} \\
&\quad - x^{(2)T}\{F_{x^{(2)}} + ... + (-1)^{p-2}F^{(p-2)}_{x^{(p)}}\} \\
&\quad - ... - x^{(p)T}F_{x^{(p)}}]\delta t|^{t_f}_{t_0} + \Psi_t \delta t]_{t_f} \\
&+[h^T\Psi_x + ... + h^{(p-1)T}\Psi_{x^{(p-1)}}]^{t_f}_{t_0} \qquad (28)
\end{aligned}$$

where $h^{(i)} = \delta x^{(i)}$ is the variation of ith derivative of x and δt is the variation of time. The necessary conditions for optimization follow by making $\delta J = 0$. This provides a set of differential equations to be satisfied by the problem and appropriate boundary conditions. From Eq. (28), it is clear that the governing differential equation is

$$F_x - F^{(1)}_{x^{(1)}} + ... + (-1)^p F^{(p)}_{x^{(p)}} = 0 \qquad (29)$$

In general, this differential equation is a $2p$th order differential equation. This equation is the extended Euler-Lagrange equation and for $p = 1$ reduces to the familiar form.

3.1.1 First Integrals

The differential equation (29) admits a number of first integrals depending on the structure of the integrand F.

- If F does not explicitly contain x, i.e., $F_x = 0$, Eq. (29) becomes

$$\frac{d}{dt}[F_{x^{(1)}} - F^{(1)}_{x^{(2)}} + ... + (-1)^{p-1}F^{(p-1)}_{x^{(p)}}] = 0 \quad (30)$$

  Hence, the optimal solution admits n first integrals $F_{x^{(1)}} - F^{(1)}_{x^{(2)}} + ... + (-1)^{p-1}F^{(p-1)}_{x^{(p)}} = K$. With a similar reasoning, if F does not explicitly contain an element of x, say x_i, correspondingly, there is a single first integral.

- If F does not explicitly depend on x and $x^{(1)}$, one can write Eq. (29) as

$$\frac{d^2}{dt^2}[F_{x^{(2)}} - F^{(1)}_{x^{(3)}} + ... + (-1)^{p-2}F^{(p-1)}_{x^{(p)}}] = 0 \quad (31)$$

  Hence, the optimal solution admits the integral $F_{x^{(2)}} - F^{(1)}_{x^{(3)}} + ... + (-1)^{p-2}F^{(p-1)}_{x^{(p)}} = K$. This argument can be extended to obtain other first integrals depending on the elements of x and their higher derivatives in the integral.

- If F does not explicitly contain t,

$$\begin{aligned}
F - x^{(1)T}\{F_{x^{(1)}} + ... + (-1)^{p-1}F^{(p-1)}_{x^{(p)}}\} \\
- x^{(2)T}\{F_{x^{(2)}} + ... + (-1)^{p-2}F^{(p-2)}_{x^{(p)}}\} \\
- ... - x^{(p)T}F_{x^{(p)}} = K \qquad (32)
\end{aligned}$$

  This property can be verified by time derivation of the left-hand and the right-hand sides. For open end time problems, where Ψ does not explicitly depend on t, i.e., $\Psi_t = 0$, the constant $K = 0$. For $p = 1$, if F is independent of time, we arrive at the familiar result $F - x^{(1)T}F_{x^{(1)}} = K$.

3.2 Higher-Order Systems

In order to address the dynamic optimization problem at hand, which consists of minimizing Eq. (2) subject to Eqs. (1) and (23), one can define F and Ψ of Eq. (28) in the following way:

$$\begin{aligned}
F(x, x^{(1)}, ..., x^{(p-1)}, x^{(p)}, u, \lambda, \xi, S, t) = L \\
+ \lambda^T(f - x^{(p)}) + \xi^T(C + S^2) \qquad (33) \\
\Psi(x(t_f), x^{(1)}(t_f), ..., x^{(p-1)}(t_f), t_f) = \Phi \quad (34)
\end{aligned}$$

where $S^2(t) = (s_1^2(t)\ s_2^2(t)\ ...\ s_r^2(t))^T$ is a $(r \times 1)$ positive slack vector for the constraints C, $\xi(t)$ are the corresponding Lagrange multipliers, $\lambda(t)$ are Lagrange multipliers corresponding to the dynamic equations. Using the general result of Eq. (28) and recognizing that F has dependence on other variables besides x and its higher derivatives, it is immediate that the optimal solution must satisfy the following differential equations:

$$\begin{aligned}
&F_x - F^{(1)}_{x^{(1)}} + ... + (-1)^p F^{(p)}_{x^{(p)}} = 0 \quad (35) \\
&F_u = 0 \quad (36) \\
&\xi_i s_i = 0,\ i = 1, ..., r \quad (37)
\end{aligned}$$

along with Eqs. (1) and (23). Using an expression for $\mathcal{H}' = L + \lambda^T f + \xi^T C$, one can easily show that Eq. (35) simplifies to

$$\mathcal{H}'_x - \mathcal{H}'^{(1)}_{x^{(1)}} + ... + (-1)^{p-1}\mathcal{H}'^{(p-1)}_{x^{(p-1)}} = (-1)^p \lambda^{(p)}, \quad (38)$$

and has a structure identical to Eq. (26). Eq. (36) can be evaluated to show that

$$F_u = \mathcal{H}'_u = 0 \qquad (39)$$

One interpretation of Eq. (39) is that on the optimal trajectory, $u(t)$ pointwise minimizes $\mathcal{H}'$. An equivalent statement is that $u(t)$ pointwise minimizes $\mathcal{H}$ subject to the constraints (23). Hence, $u^*(t)$ of Eq. (39) also satisfies the following property

$$\mathcal{H}(\mathbf{x}(t), u^*(t), \lambda, t) = \min_{u(t)\in\mathcal{A}(t,\mathbf{x})} \mathcal{H}(\mathbf{x}(t), u(t), \lambda, t). \tag{40}$$

identical to Eq. (22). Eq. (37) has the following interpretation: if $C_i < 0$, $\xi_i = 0$ and if $C_i = 0$, $\xi_i > 0$.

Eq. (38) requires boundary conditions on higher derivatives of λ. These boundary conditions can be obtained from Eq. (28). Since initial conditions on $x(t)$, $x^{(1)}(t)$,...,$x^{(p-1)}(t)$ are specified, $h(t_0)$, $h^{(1)}(t_0)$,...,$h^{(p-1)}(t_0)$ are zero. Since terminal conditions on these variables are free, the terms associated with $h(t_f)$, $h^{(1)}(t_f)$,...,$h^{(p-1)}(t_f)$ must be identically zero. This reasoning gives us the following boundary conditions:

$$\begin{aligned}&[\mathcal{H}'_{x^{(k+1)}} - \mathcal{H}'^{(1)}_{x^{(k+2)}} + ... + (-1)^{p-k-2}\mathcal{H}'^{(p-k-2)}_{x^{(p-1)}} \\ &\quad - (-1)^{p-k-1}\lambda^{(p-k-1)} + \Phi_{x^{(k)}}]_{t_f} = 0, \\ &k = 0, .., p-1\end{aligned} \tag{41}$$

On comparing Eqs. (27) with (41) and bringing these to a common index, it can be shown that the two are identical.

In summary, the functions ψ_p defined in Section 2.2 are essentially same as λ introduced in this section. Both satisfy the same differential Eqs. (26) or (38) and the same boundary conditions (27) or (41). The optimal $u(t)$ in both approaches are the same.

3.2.1 First Integrals

For the system under consideration, $F = \mathcal{H}' - \lambda^T x^{(p)}$ and $\mathcal{H}' = L + \lambda^T f + \xi^T C$. If L, f, and C are not explicit functions of time, according to Section 3.1.1, the solution has a first integral given by Eq. (32). On siplifying, it can be shown that this equation reduces to

$$\begin{aligned}&\mathcal{H}' - x^{(1)^T}[\mathcal{H}'_{x^{(1)}} - \mathcal{H}'^{(1)}_{x^{(2)}} + ... + (-1)^{p-2}\mathcal{H}'^{(p-2)}_{x^{(p-1)}} \\ &\qquad + (-1)^p\lambda^{(p-1)}] \\ &-x^{(2)^T}[\mathcal{H}'_{x^{(2)}} - \mathcal{H}'^{(1)}_{x^{(3)}} + ... + (-1)^{p-3}\mathcal{H}'^{(p-3)}_{x^{(p-1)}} \\ &\qquad + (-1)^{p-1}\lambda^{(p-2)}] - ... \\ &-x^{(p-1)^T}[\mathcal{H}'_{x^{(p-1)}} + \lambda^{(1)}] = K\end{aligned} \tag{42}$$

With the equivalence between λ and ψ_p of Section 2.2, one can rewrite Eq. (42) using Eq. (27) as

$$\mathcal{H}' + \psi_1^T x^{(1)} + \psi_2^T x^{(2)} + ... + \psi_{p-1}^T x^{(p-1)} = K \tag{43}$$

3.3 First-Order Representation

An alternative approach to derive the necessary conditions for minimum of the problem posed in Section 1 is to first express the higher-order differential Eqs. (1) as a set of first-order differential equations and then to use the classical results of the minimum principle. This section outlines this procedure and it is demonstrated that the results derived in this way are consistent with the results from the earlier approaches described in the paper where the system equations are left in the higher-order form.

A simple way to express the higher-order system by first-order differential equations is through the definition of the extended vector $\mathbf{x}(t)^T = (x(t)^T\ x^{(1)}(t)^T\ ...\ x^{(p-1)}(t)^T)^T$ that was introduced in Section 2. This vector has a dimension of $(np \times 1)$. Let the individual $(n \times 1)$ subvectors of $\mathbf{x}(t)$ be renamed as $\mathbf{x}_1(t) = x(t)$, $\mathbf{x}_2(t) = x^{(1)}(t)$, ...,$\mathbf{x}_p(t) = x^{(p-1)}(t)$. With this definition,

$$\begin{aligned}\dot{\mathbf{x}}_1(t) &= \mathbf{x}_2(t) \\ \dot{\mathbf{x}}_2(t) &= \mathbf{x}_3(t) \\ &\vdots \\ \dot{\mathbf{x}}_{p-1}(t) &= \mathbf{x}_p(t) \\ \dot{\mathbf{x}}_p(t) &= f(\mathbf{x}_1(t), \mathbf{x}_2(t), ..., \mathbf{x}_p(t), u(t), t)\end{aligned} \tag{44}$$

The cost functional J can also be restated using the definitions of $\mathbf{x}_i(t)$. The result of Section 3.1 can now be used to derive the necessary conditions of optimum for this new problem as long as we use $p = 1$ in Eq. (28).

With the dynamic optimization problem expressed in first-order form, the definitions of F and Ψ in Eq. (28) are:

$$\begin{aligned}&F(\mathbf{x}_1(t), ..., \mathbf{x}_p(t), \dot{\mathbf{x}}_1(t), ...\dot{\mathbf{x}}_p(t), \lambda_1, ..., \lambda_p, u(t), \\ &\xi, S, t) = L + \lambda_1^T(\mathbf{x}_2 - \dot{\mathbf{x}}_1) + ... \\ &+\lambda_{p-1}^T(\mathbf{x}_p - \dot{\mathbf{x}}_{p-1}) + \lambda_p^T(f - \dot{\mathbf{x}}_p) + \xi^T(C + S^2)\end{aligned} \tag{45}$$

$$\Psi(\mathbf{x}_1(t_f), \mathbf{x}_2(t_f), ..., \mathbf{x}_p(t_f), t_f) = \Phi \tag{46}$$

where $S^2(t) = (s_1^2(t)\ s_2^2(t)\ ...\ s_r^2(t))^T$ is a $(r \times 1)$ positive slack vector for the constraints C, $\xi(t)$ are the corresponding Lagrange multipliers, and $\lambda_i(t)$ are Lagrange multipliers corresponding to the state equations.

Using the result of Eq. (28) with $p = 1$, it is immediate that the optimal solution must satisfy the following differential equations:

$$F_{\mathbf{x}_i} - F^{(1)}_{\dot{\mathbf{x}}_i} = 0,\ i = 1, ..., p \tag{47}$$

$$F_u = 0 \tag{48}$$

$$\xi_i s_i = 0,\ i = 1, ..., r \tag{49}$$

along with Eqs. (23), (44) and (46). For this problem, we can define two different Hamiltonians and it is important to distinguish between the two. For this functional given in Eq. (45), the classical definition of augmented Hamiltonian is $\mathcal{H}'_c = L + \lambda_p^T f + \xi^T C + \lambda_1^T \mathbf{x}_2 + ... + \lambda_{p-1}^T \mathbf{x}_p$. Here, the superscript c has been used to distinguish it from the Hamiltonian $\mathcal{H}' = L + \lambda_p^T f + \xi^T C$, described similar to Sections 2.1.1 and 3.2.

One can now show that Eq. (47) simplifies to

$$\begin{aligned} &\dot{\lambda}_1 + \mathcal{H}'_{c\mathbf{x}_1} = 0 \\ &\dot{\lambda}_2 + \mathcal{H}'_{c\mathbf{x}_2} = 0 \\ &\vdots \\ &\dot{\lambda}_{p-1} + \mathcal{H}'_{c\mathbf{x}_{p-1}} = 0 \\ &\dot{\lambda}_p + \mathcal{H}'_{c\mathbf{x}_p} = 0 \end{aligned} \tag{50}$$

or alternatively to

$$\begin{aligned} &\dot{\lambda}_1 + \mathcal{H}'_{\mathbf{x}_1} = 0 \\ &\dot{\lambda}_2 + \lambda_1 + \mathcal{H}'_{\mathbf{x}_2} = 0 \\ &\vdots \\ &\dot{\lambda}_{p-1} + \lambda_{p-2} + \mathcal{H}'_{\mathbf{x}_{p-1}} = 0 \\ &\dot{\lambda}_p + \lambda_{p-1} + \mathcal{H}'_{\mathbf{x}_p} = 0 \end{aligned} \tag{51}$$

Since Eq. (51) has a structure similar to Eq. (18), we prefer to use $\mathcal{H}'$ in our work instead of $\mathcal{H}'_c$. On eliminating λ_1 through λ_{p-1} in this equation, similar to Section 2.1.1, one can show that

$$\mathcal{H}'_x - \mathcal{H}'^{(1)}_{x^{(1)}} + ... + (-1)^{p-1} \mathcal{H}'^{(p-1)}_{x^{(p-1)}} = (-1)^p \lambda_p^{(p)}, \tag{52}$$

Eq. (48) can be evaluated to show that

$$F_u = \mathcal{H}'_u = 0 \tag{53}$$

which can also be interpreted as finding $u^*(t)$ such that

$$\mathcal{H}(\mathbf{x}(t), u^*(t), \lambda, t) = \min_{u(t) \in \mathcal{A}(t,\mathbf{x})} \mathcal{H}(\mathbf{x}(t), u(t), \lambda, t). \tag{54}$$

Eq. (49) has the following interpretation: if $C_i < 0$, $\xi_i = 0$ and if $C_i = 0$, $\xi_i > 0$. Eq. (52) requires boundary conditions on λ. These boundary conditions at t_f are obtained by a logic similar to Section 3.2 and it can be shown that they are consistent with those obtained in the other approaches.

From these steps, it is now clear that a λ_p in this section is same as λ of Section 3.2 as well as ψ_p of Section 2.1.1. Also, a λ_k, $k = 1, ..., p-1$ defined in this section is the same as ψ_k of Section 2.1.1. The definition of $\mathcal{H}'$ is consistent within all approaches and the optimal $u(t)$ is obtained by $\min_{u(t) \in \mathcal{A}(t,\mathbf{x})} \mathcal{H}(\mathbf{x}(t), u(t), \lambda, t)$.

3.3.1 First Integrals

For the system under consideration, $F = \mathcal{H}' + \lambda_1^T(\mathbf{x}_2 - \dot{\mathbf{x}}_1) + ... + \lambda_{p-1}^T(\mathbf{x}_p - \dot{\mathbf{x}}_{p-1}) - \lambda_p^T \dot{\mathbf{x}}_p$, where $\mathcal{H}' = L + \lambda_p^T f + \xi^T C$. According to the result of Section 3.1.1, if F is not an explicit function of time, the solution has a first integral given by Eq. (32) simplified for $p = 1$. This condition is

$$F - \sum_{i=1}^{p} \dot{\mathbf{x}}_i F_{\dot{\mathbf{x}}_i} = K \tag{55}$$

By evaluation, it can be shown equivalent to

$$\mathcal{H}' + \lambda_1^T \mathbf{x}_2 + ... + \lambda_{p-1}^T \mathbf{x}_p = K \tag{56}$$

This result is consistent with the result of Eq. (43). Further, this is also equivalent to the statement $\mathcal{H}'_c = K$, consistent with the classical first-order result.

4 Example

From Sections 2 and 3, it is clear that there is a complete equivalence between the optimization results whether they are derived using the higher-order or the first-order form. This equivalence holds for all systems, linear or nonlinear, as long as the constraints are consistent with the class proposed in this paper. In this section, we provide an example of a nonlinear system which can be written both in the first-order and higher-order form.

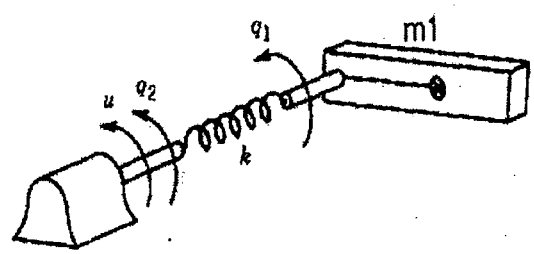

Figure 1: Single-link robot with joint flexibility.

The example is of a single link manipulator rotating in a vertical plane driven through a flexible drive train [21], shown in Fig. 1. The system has two degrees-of-freedom and the equations of motion are

$$I q_1^{(2)} + mgl \sin q_1 + k(q_1 - q_2) = 0 \tag{57}$$

$$J q_2^{(2)} - k(q_1 - q_2) = u, \tag{58}$$

where I and J are respectively the link and actuator moment of inertia, m is the mass of link with mass center at a distance l from the joint, k is the stiffness of the drive train, g is the gravity constant, and u is the actuator torque. Let the objective be to steer the system from a given set of initial conditions on q_1, q_2, $q_1^{(1)}$, and $q_2^{(1)}$ at t_0 to an unspecified goal point while minimizing a cost $J = \int_{t_0}^{t_f} u^2 dt$. The trajectory must satisfy the constraint $-1 \leq u \leq 1$ during motion.

For the purposes of optimization, one can proceed in several alternative ways. For brevity, we will describe only two out of four alternative ways in detail. The two second-order differential equations describing the system have a special structure. From Eq. (57), q_2 can be written explictly in terms of q_1 and its second derivative. On substituting this expression of q_2 in Eq. (58), we obtain a single fourth-order differential equation in the variable q_1 up to its fourth derivatives

$$q_1^{(4)} = \alpha_1 u + (\alpha_2 \cos q_1 + \alpha_3) q_1^{(2)} + \alpha_4 \sin q_1 {q_1^{(1)}}^2 + \alpha_5 \sin q_1, \quad (59)$$

where α_i are constants with $\alpha_1 = \frac{K}{IJ}$, $\alpha_2 = -\frac{mgl}{I}$, $\alpha_3 = -\frac{K(I+J)}{IJ}$, $\alpha_4 = \frac{mgl}{I}$, and $\alpha_5 = -\frac{mgkl}{IJ}$. This differential equation has the structure of Eq. (1) with $n = 1$ and $m = 1$.

Also, the fourth-order differential equation of Eq. (59) could be written in the first-order form

$$\begin{aligned} x_1^{(1)} &= x_2 \\ x_2^{(1)} &= x_3 \\ x_3^{(1)} &= x_4 \\ x_4^{(1)} &= \alpha_1 u + (\alpha_2 \cos x_1 + \alpha_3) x_1^{(2)} + \sin x_1 {x_2}^2, \\ &\quad + \alpha_5 \sin x_1 \end{aligned} \quad (60)$$

where $x_1 = q_1$. This differential equation has the structure of Eq. (1) with $n = 4$ and $m = 1$.

Eqs. (57) and (58) are both second-order differential equations. On solving for $q_1^{(2)}$ and $q_2^{(2)}$ from these, one obtains the equations to have the form Eq. (1) with $n = 2$ and $m = 1$. The fourth possibility is that Eqs. (57) and (58) are each reduced to two first-order equations, thereby, producing a form of Eq. (1) with $n = 4$ and $m = 1$.

The optimization of this problem could be addressed using any of the four alternative descriptions of the system. For brevity, we will only show the equivalence of the results for the first two forms, i.e., the fourth-order form of Eq. (59) and the first-order form of Eq. (60).

4.1 Fourth-Order Form

Using the fourth-order form, the $\mathcal{H}' = \mathcal{H} = u^2 + \lambda[\alpha_1 u + (\alpha_2 \cos q_1 + \alpha_3) q_1^{(2)} + \alpha_4 \sin q_1 {q_1^{(1)}}^2 + \alpha_5 \sin q_1]$. The feasible control is defined in the set $\mathcal{U} = [-1, 1]$. From the extended form of Pontryagin's principle, we need to select u such that $\mathcal{H}'$ is minimized. Since $u^2 + \lambda\alpha_1 u$ is the term in $\mathcal{H}'$ that is dependent on u, it can also be written as $(u + \frac{\lambda\alpha_1}{2})^2 - (\frac{\lambda\alpha_1}{2})^2$. From this form, it is evident that u^*, the minimum of u within $\mathcal{U}$ satisfies the following switching structure:

$$u^* = \left\{ \begin{array}{rr} 1, & \alpha_1\lambda < -2 \\ -\lambda\alpha_1/2, & -2 \leq \alpha_1\lambda \leq 2 \\ -1, & \alpha_1\lambda > 2 \end{array} \right\} \quad (61)$$

On evaluating Eq. (38), the following fourth-order differential equation in λ results

$$\lambda^{(4)} - \lambda^{(2)}(\alpha_3 + \alpha_2 \cos q_1) - \lambda\alpha_5 \cos q_1 = 0 \quad (62)$$

In summary, the optimal solution is characterized by the fourth-order differential Eq. (59), the fourth-order Lagrange multipler Eq. (62), and the switching structure of Eq. (61).

4.2 First-Order Form

One can either work with $\mathcal{H}_c'$ and use Eq. (50) to determine the Lagrange mutiplier equation or use $\mathcal{H}'$ with Eq. (51) to determine the Lagrange multipliers. The two expressions are: $\mathcal{H}' = \mathcal{H} = u^2 + \lambda_4[\alpha_1 u + (\alpha_2 \cos x_1 + \alpha_3) x_1^{(2)} + \sin x_1 {x_2}^2 + \alpha_5 \sin x_1]$ and $\mathcal{H}_c' = \mathcal{H}_c = u^2 + \lambda_1 x_2 + \lambda_2 x_3 + \lambda_3 x_4 + \lambda_4[\alpha_1 u + (\alpha_2 \cos x_1 + \alpha_3) x_1^{(2)} + \sin x_1 {x_2}^2 + \alpha_5 \sin x_1]$. Using the results of Section 3.3, it can be shown that the Lagrange multiplier equations in both cases are

$$\lambda_4^{(4)} - \lambda_4^{(2)}(\alpha_3 + \alpha_2 \cos x_1) - \lambda_4\alpha_5 \cos x_1 = 0 \quad (63)$$

With either expression of the Hamiltonian, it is clear that u^*, the minimum of u within $\mathcal{U}$ satisfies the following switching structure:

$$u^* = \left\{ \begin{array}{rr} 1, & \alpha_1\lambda_4 < -2 \\ -\lambda_4\alpha_1/2, & -2 \leq \alpha_1\lambda_4 \leq 2 \\ -1, & \alpha_1\lambda_4 > 2 \end{array} \right\} \quad (64)$$

As expected, Eqs. (62) and (63) are identical as long as x_1 is interpreted as q_1 and λ is interpreted as λ_4. Similarly, Eqs. (61) and (64) are identical.

5 Numeric Computation

Even though the focus of this paper is on the underlying theory of optimization of higher-order systems,

some brief insights into the numerical aspects of the solution procedure using the higher-order form will be discussed. We will outline a scheme for solving Eqs.(1) and (38) with the switching structure given by (40). The procedure will be compared with an equivalent procedure applied to first order form of the optimality equations.

Our approach for solution is based on theory of collocation and consists of the following steps: (i) choose N intermediate points (nodes) between t_0 and t_f and divide the total time into $N+1$ intervals; (ii) choose an admissible form for $q(t)$, $\lambda(t)$, and $u(t)$ in each of these intervals, i.e., form a base solution and add mode functions to refine the solution; (iii) impose continuity of $q(t)$ and $\lambda(t)$ upto $p-1$th derivative at the N intermediate node points but allow discontinuity of $u(t)$; (iv) solve for the base solution using the given boundary conditions of q and λ at t_0 and t_f and continuity at the node points; (v) satisfy the higher-order differential equations by choosing collocation points within each interval, i.e., each collocation point provides n equations corresponding to Eqs.(1) and another n to Eq. (38); (vi) pose this problem as a nonlinear programming problem with the objective to find the best mode coefficients and the location of the nodes such that $\mathcal{H}'$ is minimized with respect to the input. The salient feature to note here is that each collocation point gives $2n$ nonlinear equations.

If the same collocation procedure was applied to the system equations and Lagrange multiplier equations, each in the first-order form, every colocation point will result in $2np$ nonlinear equations in the nonlinear programming problem. Most computer implementations of the classical first-order form routinely implement this. As is evident, if $p = 1$, the computations with the higher-order form and the first-order form are identical but as p becomes larger, the higher-order approach provides a lot fewer equations in the nonlinear programming problem. For the example problem, the higher-order forms of Eqs.(59) and (62) have $p = 4$. Therefore, it is more computationally efficient to use the higher-order solution scheme outlined in this section. Currently, a general purpose program is being developed based on these ideas.

6 Conclusion

In this paper, dynamic optimization of systems described by higher-order differential equations is addressed. The minimum principle for higher-order systems is derived directly from their higher-order forms and the results are confirmed by classical theory using first-order forms. The optimality conditions are derived using both Hamilton-Jacobi theory, thereby, extending Pontryagin's theory and through variational calculus. It is shown that the different approaches lead to the same results. The result applicable to higher-order systems is illustrated by an example. Also, insights are provided why numerical implementation of the higher-order formulation is computationally more efficient compared to the classical first-order approach. Currently, general purpose programs are being developed that exploit the features of this higher-order approach.

Acknowledgments The author acknowledges the support of Thailand Toray Science Foundation.

References

[1] Agrawal, S. K. and Veeraklaew, T., A Higher-Order Method for Dynamic Optimization of a Class of Linear Systems, *Journal of Dynamic Systems, Measurement, and Control, Transactions of the ASME*, Vol. 118, Dec. 1996, 786-791.

[2] Agrawal, S. K. and Xu, X., A Higher-Order Method for Dynamic Optimization of a Class of Linear Time-Varying Dynamic Systems, *Journal of Vibrations and Control*, Vol. 4, No. 4, 1997, 379-396.

[3] Agrawal, S. K., Claewplodtook, P., and Fabien, B. C., Optimal Trajectories of Open-Chain Mechanical Systems: A New Solution Without Lagrange Multipliers, *Journal of Dynamic Systems, Measurement, and Control, Transactions of the ASME*, Vol. 120, No. 1, 1998, 134-136.

[4] Agrawal, S. K. and Faiz, N., *Journal of Optimization Theory and Applications*, 1998.

[5] Agrawal, S.K. and Fabien, B.C., *Optimization of Dynamic Systems*, Kluwer Academic Publishers, to appear 1999.

[6] Berkovitz, L. D., Variational Methods in Problems of Control and Programming, *Journal of Mathematical Analysis and Applications*, Vol. 3, 1961, 145-169.

[7] Bliss, G. A., *Lectures on the Calculus of Variations*, The University of Chicago Press, Chicago, 1946.

[8] Bolza, O., *Lectures on the Calculus of Variations*, G. E. Stechert and Company, 1931.

[9] Bryson, A. E and Ho, Y.C., *Applied Optimal Control*, Hemisphere Publishing Company, 1975.

[10] Fliess, M., Levine, J., Martin, P., Rouchon, P., Flatness and Defect of Nonlinear Systems: Introductory Theory and Examples, *Journal of Control*, Vol. 61, No. 2, 1995, 1327-1361.

[11] Forsyth, A. R., *Calculus of Variations*, Dover Publications, Inc., New York, 1960.

[12] Gelfand, I.S. and Fomin, S.V., *Calculus of Variations*, Prentice Hall Book Company, 1963.

[13] Isidori, A., *Nonlinear Control Systems*, Springer Verlag, 1994.

[14] Kirk, D.E., *Optimal Control Theory: An Introduction*, Prentice Hall Electrical Engineering Series, 1970.

[15] Leitmann, G., *Optimization Techniques*, Academic Press Inc., London, 1962.

[16] Murray, R.R., Li, Z.,and Sastry, S.S., *A Mathematical Introduction to Robotic Manipulation*, CRC Press, 1993.

[17] Nieuwstadt, M. J. and Murray, R. M., Real Time Trajectory Generation for Differentially Flat Systems, *Int'l. J. Robust and Nonlinear Control*, 1998.

[18] Pontryagin, L.S., Boltyanskii, V. G., Gamkrelidze, R. V., and Mishchenko, E. F., *The Mathematical Theory of Optimal Process*, John Wiley and Sons Inc., New York, 1962.

[19] Sage, A.P.,*Optimum Systems Control*, Prentice-Hall Inc., Englewood Cliffs, N.J., 1968.

[20] Schlemmer, M. and Agrawal, S. K., Globally Feedback Linearizable Time-Invariant Systems: Optimal Solution Using Mayer's Cost, submitted for publication *Journal of Dynamic Systems, Measurement, and Control, Transactions of the ASME*, 1998.

[21] Spong, M.W., Vidyasagar M., *Robot dynamics and control*, John Wiley and Sons, 1986.