JOURNAL OF GUIDANCE, CONTROL, AND DYNAMICS
Vol. 29, No. 4, July–August 2006

Optimal Finite-Time Feedback Controllers for Nonlinear Systems with Terminal Constraints

Srinivas R. Vadali* and Rajnish Sharma†
Texas A&M University, College Station, Texas 77843-3141

This paper presents a dynamic-programming approach for determining finite-time optimal feedback controllers for nonlinear systems with nonlinear terminal constraints. The method utilizes a polynomial series expansion of the cost-to-go function with time-dependent gains that operate on the state variables and constraint Lagrange multipliers. These gains are computed from backward integration of differential equations with terminal boundary conditions, derived from the constraint specifications. The differential equations for the gains are independent of the states. The Lagrange multipliers at any particular time are evaluated from the knowledge of the current state and the gain values. Several numerical examples are considered to demonstrate the applications of this methodology. The accuracy of the method is ascertained by comparing the results with those obtained by using open-loop solutions to the respective problems. Finally, results of the application of the developed methodology to a spacecraft detumbling problem are presented.

Introduction

SOLUTIONS to optimal control problems can be obtained by solving the Hamilton–Jacobi–Bellman (HJB) equation, a nonlinear partial differential equation. The method of characteristics, which is typically implemented via shooting algorithms, results in open-loop solutions to the HJB equation. The linear-quadratic (LQ) problem is a special case for which there is an exact, closed-form feedback solution for the HJB equation, with the control gains resulting from the solution to a Riccati equation and a set of linear auxiliary differential equations. Other special cases of norm-invariant nonlinear systems exist for which the optimal control involves a linear feedback of the states. An example of this type of a problem is that of the optimal detumbling of a rigid body using continuous controls, which minimize a quadratic, infinite-horizon performance index, based on the kinetic energy.[1] Many approaches have been devised in the controls literature to obtain approximate solutions to the unconstrained HJB equation, in general settings. In this paper, attention is focused on problems for which the HJB equation admits smooth solutions.

Power-series solution methods to the infinite-horizon problems can be traced back to the work of A'lbrekht[2] who used Lyapunov functions to construct optimal controls for nonlinear systems, in feedback form. The convergence proof is based on the asymptotic stability of the linear part of the closed-loop system. Lukes[3] formalized the power-series approach by expanding the cost and controls about the origin. Dabbous and Ahmed[4] showed the application of the series solution method to the problem of nonlinear optimal angular momentum regulation of a satellite. Carrington and Junkins[5] used the series expansion of the costate vector to solve the optimal finite-horizon, spacecraft attitude regulation problem in feedback form. Their approach is different from the series solution of the HJB equation in which the scalar cost function is expanded in a power series. The control law obtained by Carrington and Junkins[5] involves nonlinear feedback of angular velocities and Euler parameters, using time-varying, matrix, and higher-order tensor gain elements. They incorporated a change of variables involving the terminal boundary conditions in order to convert a general maneuver problem into a zero-set-point problem. In addition, they chose a long enough horizon for completing the maneuver, thus requiring only the methodology of the asymptotically stable, infinite-horizon regulator.

Yoshida and Loparo[6] presented an approximation theory for the optimal regulation of nonlinear systems with additive controls using a lexicographic listing of vectors containing monomials of different degrees. This arrangement allows the use of Carleman linearization for the theoretical analysis of convergence. Extension of the series solution methodology to the solution of the Hamilton–Jacobi–Isaacs equation can be found in Huang and Lin[7] and Tsiotras et al.[8]

Approximations of other forms can also be found in the literature on suboptimal control. Garrard et al.[9] presented an approximation method by neglecting the time derivatives of certain nonlinear terms in the HJB equation. One can establish a connection between this work and the state-dependent Riccati equation approach.[10] Tewari[11] presented an iterative approximation technique for the spacecraft attitude control problem, which uses a specific form of a positive-definite cost function (Lyapunov function) chosen a priori. Extension of this technique to tracking maneuvers can be found in Sharma and Tewari.[12]

Cimen and Banks[13] introduced an "approximating sequence of Riccati equations" approach to nonlinear optimal control problems with nonquadratic performance indices. However, they did not consider terminal constraints. In this approach, the optimal control is obtained by solving a sequence of time-varying Riccati differential equations until convergence is achieved. The result is an optimal control law for the original nonlinear system, which has a linear state feedback structure, with time-varying gains. However, the feedback control solution has to be recomputed for each new initial condition. The $\theta - D$ suboptimal approach of Xin et al.[14] also results in an approximate solution to the HJB equation. It involves the addition of small perturbations to the original cost function such that a closed-form solution to the control can be achieved. This method can also be classified as an inverse optimal control approach.

Inverse optimal control approaches obtain a performance index that is minimized, given a stabilizing feedback control law. However, often, the performance index might not have known physical significance. Examples of the applications of this technique can be found in Vadali and Junkins[15] and Kristic and Tsiotras.[16]

Beard et al.[17] developed a Galerkin successive approximation technique to solve the generalized HJB (GHJB) equation, a linear partial differential equation. The sequence of solutions converges to the solution of the HJB equation in the limit. Extensions of the

Presented as AAS Paper 05-182 at the AAS/AIAA 15th Space Flight Mechanics Conference, Copper Mountain, CO, 23–27 January 2005; received 24 March 2005; revision received 20 July 2005; accepted for publication 25 July 2005.

*Stewart & Stevenson Professor, Department of Aerospace Engineering, 3141 TAMU; svadali@aero.tamu.edu. Associate Fellow AIAA.

†Graduate Student, Department of Aerospace Engineering, 3141 TAMU; aeroraj@neo.tamu.edu. Student Member AIAA.

methodology to solve robust optimal control problems can be found in Beard and McLain.[18] Other approaches exploit the fact that the optimal control at any time is dependent on the local gradient of the cost-to-go. Hence a local approximation procedure is also possible to approximate the cost function in the neighborhood of the current state.[19] Park and Tsiotras[20,21] use the wavelet basis functions to solve the GHJB equation using the iterative Galerkin procedure as well as the collocation approach.

As just discussed, much of the work has been focused on the infinite-time regulation problem, but the fixed final-time problem with nonlinear terminal constraints has not received its due attention. The primary reason is the need to deal with the hard terminal constraints and the associated computation of the constraint Lagrange multiplier. A special case of the hard terminal constraint problem is one where the final states are specified a priori or can be solved for uniquely from the constraint equation. For such problems, in a series of papers, Guibout and Scheeres[22] and Park and Scheeres[23,24] proposed an alternate approach based on canonical transformations and the Taylor's series expansions of the Hamiltonian and generating functions. They have demonstrated a computational technique to solve in closed form a class of two-point boundary-value problems arising in the context of impulsive orbit-transfer problems, formation flying, as well as optimal control problems. In this approach, the state and costate vectors at a particular time instant are viewed as a result of a canonical transformation from the initial state and costate vectors. Different types of generating functions can be chosen depending on the class of problems (boundary conditions) being solved. A generating function of one type can be converted into that of another type via the Legendre transformation. Reference 24 contains applications of the HJB theory to the minimum-time control problem of the double integrator as well as a singular control problem for a linear system with a quadratic cost function. Explicit nonlinear terminal constraints have not been accounted for in Refs. 22–24.

The sweep method[25] was developed for solving finite-time LQ problems with linear terminal constraints. The method is based on two steps. As presented in Ref. 25, the first step expresses the costate vector at any time instant as a linear combination of the current state vector and the constraint Lagrange multiplier vector. The second step in the process requires that the terminal constraint also be expressed as a linear combination of the same form, as already mentioned. A symmetry property, involving the gains of the two linear combinations, is obtained after the derivation process for the differential equations governing the evolution of the gain matrices is completed. Similarity between the sweep method and dynamic programming for a class of LQ problems is shown in Bryson.[26] An alternate derivation using the generating function approach for solving the HJB equation is also presented in Park and Scheeres[23] for the solution to the LQ problem. The terminal constraint is a given final state vector, and a constraint Lagrange multiplier vector is not required in this approach. The HJB approach is more direct and natural than the sweep method.

In this paper, the dynamic-programming approach, based on power-series expansion of the cost function, is extended to deal with finite-time feedback control of nonlinear systems with nonlinear terminal constraints. In so doing, a compact derivation of the results of the sweep method for LQ problems is obtained as a special case. The power-series solution methodology is used, but, unlike in the works just discussed, explicit use is made of the constraint Lagrange multipliers. The methodology is illustrated first on a problem that has an analytical solution. Subsequently, a two-dimensional problem with a nonlinear terminal constraint is treated. Finally, an example of optimal detumbling of a spacecraft is presented.

Statement of the Problem

The main objective in this study is the determination of a nonlinear feedback control law to minimize a general Bolza performance index (cost) defined as

$$J = \phi[x(t_f)] + \int_{t_0}^{t_f} L(x, u, \tau)\,\mathrm{d}\tau \tag{1}$$

for a nonlinear dynamical system:

$$\dot{x} = f(x, u, t) \tag{2}$$

where $f: \Re^{n+m+1} \rightarrow \Re^n$ is a smooth, analytic, vector-valued function with $x \in \Re^n$ and $u \in \Re^m$. The initial time t_0, the final time t_f, and the initial condition on the state vector x_0 are assumed to be given. It is also assumed that $f(0, 0, t) = 0$ and the control vector u is unconstrained. The terminal constraint is given by $\psi[x(t_f)] - \psi_f = 0$; $\psi, \psi_f \in \Re^{p \leq n}$, where ψ_f, the value of the constraint, is defined to satisfy the condition $\psi(0) = \psi_f$. Equation (1), when written with the current time as the lower limit of integration, represents the cost-to-go function, and it is termed the value function in the dynamic-programming literature. It is also assumed that the value function and the controls are smooth with respect to the states, Lagrange multipliers, and time.

Derivation of the Important Results

Herein, the variational approach to the derivation of the necessary conditions is briefly outlined first, and then the results are related to those obtained from dynamic programming.

The terminal constraint and Eq. (2) are augmented to the cost function, and the augmented cost function is written as

$$J_a = \phi[x(t_f)] + \boldsymbol{\nu}^T\{\psi[x(t_f)] - \psi_f\} + \int_{t_0}^{t_f} \{L(x, u, t) + \boldsymbol{\lambda}^T[f(x, u, t) - \dot{x}]\}\,\mathrm{d}t \tag{3}$$

where $\boldsymbol{\nu}$ is a vector of constant Lagrange multipliers that enforce the terminal constraints and $\boldsymbol{\lambda}$ is the costate vector. The necessary conditions for optimality can be derived by using the Hamiltonian formalism, which begins with the definition of the Hamiltonian:

$$H = L(x, u, t) + \boldsymbol{\lambda}^T f(x, u, t) \tag{4}$$

The first variation of the augmented cost function[25] is written as

$$\delta J_a = \left[\frac{\partial\phi}{\partial x(t_f)} + \frac{\partial\psi}{\partial x(t_f)}^T \boldsymbol{\nu} - \boldsymbol{\lambda}(t_f)\right]^T \delta x(t_f) + \{\psi[x(t_f)] - \psi_f\}^T \mathrm{d}\boldsymbol{\nu} + \int_{t_0}^{t_f}\left[\left(\frac{\partial H}{\partial x} + \dot{\boldsymbol{\lambda}}\right)^T \delta x + \left(\frac{\partial H}{\partial \boldsymbol{\lambda}} - \dot{x}\right)^T \delta\boldsymbol{\lambda} + \left(\frac{\partial H}{\partial u}\right)^T \delta u\right]\mathrm{d}t \tag{5}$$

First-order necessary conditions that must be satisfied by the optimal state and control are obtained by requiring that the first variation δJ_a in Eq. (5) vanish, for admissible variations in x, u, λ, and ν. The necessary conditions for extremizing the augmented performance index are[25]

$$\dot{x} = \frac{\partial H}{\partial \boldsymbol{\lambda}}, \qquad \dot{\boldsymbol{\lambda}} = -\frac{\partial H}{\partial x} \tag{6}$$

$$u^*(t) = \min_u(H) \tag{7}$$

where $u^*(t)$ is the optimal control.

The transversality condition is

$$\boldsymbol{\lambda}(t_f) = \frac{\partial\phi[x(t_f)]}{\partial x(t_f)} + \left\{\frac{\partial\psi[x(t_f)]}{\partial x(t_f)}\right\}^T \boldsymbol{\nu} \tag{8}$$

A further necessary condition is that the terminal constraint must be satisfied.

Dynamic-programming formalism also provides necessary and sufficient conditions for the preceding minimization problem in the form of the HJB equation and the associated boundary conditions. Explicit feedback control laws can be obtained from the solutions to the HJB equation, a partial differential equation for the optimal value function (cost-to-go) J^* as shown next[25]:

$$\frac{\partial J^*}{\partial t} = -H\left[x(t), u^*(t), \frac{\partial J^*}{\partial x}, t\right] \tag{9}$$

with the boundary condition

$$J^*[x(t_f), t_f] = \phi[x(t_f)];\ x(t_f) \quad \text{such that} \quad \psi[x(t_f)] = \psi_f \tag{10}$$

The key result that connects dynamic programming and the variational approach is the relationship between the costate vector and the gradient of the value function[25,26]:

$$\boldsymbol{\lambda}^*(t) = \frac{\partial J^*[x(t), t]}{\partial x(t)} \tag{11}$$

Indeed, using the Lagrange multiplier rule, Eq. (10) can also be written as

$$J^*[x(t_f), t_f] = \phi[x(t_f)] + \boldsymbol{\nu}^T\{\psi[x(t_f)] - \psi_f\} \tag{12}$$

and the partial derivative of the preceding expression with respect to $x(t_f)$ results in the boundary condition of Eq. (8). A detailed treatment of the boundary conditions can be found in Dreyfus.[27]

Another key sensitivity result can be arrived at by considering the effect of the differential in the constraint Lagrange multiplier vector $\mathrm{d}\boldsymbol{\nu}$ on the cost, in the neighborhood of the optimal solution. It can be observed from Eq. (5) that if all of the necessary conditions are satisfied then the gradient with respect to $\boldsymbol{\nu}$, of the augmented cost function, along the optimal trajectory is

$$\frac{\partial J_a}{\partial \boldsymbol{\nu}} = \frac{\partial J[x(t_0), t_0]}{\partial \boldsymbol{\nu}} + \{\psi[x(t_f)] - \psi_f\} = 0 \tag{13}$$

Because the terminal conditions are also satisfied as part of the necessary conditions and because the initial time is arbitrary, the preceding result can also be stated as

$$\frac{\partial J^*[x(t), t]}{\partial \boldsymbol{\nu}} = 0 \tag{14}$$

The partial derivative of Eq. (12) with respect to $\boldsymbol{\nu}$ shows that the preceding condition is also satisfied at the final time:

$$\frac{\partial J^*[x(t_f), t_f]}{\partial \boldsymbol{\nu}} = \{\psi[x(t_f)] - \psi_f\} = 0 \tag{15}$$

Equation (14) is the main result that will be exploited in this paper. This important result has not been dealt with in any of the standard references found in the literature. It is also important to point out that unlike the solution obtained from a two-point boundary-value problem a solution that satisfies the HJB equation, if it exists, satisfies both the necessary and sufficient conditions for a minimum. In the next section, the dynamic-programming approach is applied to the fixed-horizon, terminally constrained, LQ problem. Interested readers can compare the derivation process with that presented in Bryson and Ho.[25]

Alternative Derivation of the LQ Terminal Controller

Consider the following LQ problem with known initial and final times:

Minimize:

$$J = \frac{1}{2}\int_{t_0}^{t_f} (x^T Q x + u^T R u)\,\mathrm{d}t \tag{16}$$

Subject to:

$$\dot{x} = Ax + Bu \tag{17}$$

and a linear terminal constraint

$$Cx(t_f) = \psi_f \tag{18}$$

The weight matrix Q is assumed to be symmetric and positive semidefinite. The control-weighting matrix R is assumed to be symmetric and positive definite. Controllability of the pair (A, B) and observability of the pair $(A, \sqrt{Q})$ are also assumed. When the number of constraints is equal to the number of states and the matrix C in Eq. (18) is invertible, then the final state can be explicitly solved for. For cases where the number of constraints is less than the dimension of the state vector, the final state is determined as a result of the optimization process.

The Hamiltonian for the problem is

$$H = \tfrac{1}{2}[x^T Q x + u^T R u] + \boldsymbol{\lambda}^T[Ax + Bu] \tag{19}$$

The application of the optimality condition results in the following relationship:

$$u = -R^{-1}B^T\boldsymbol{\lambda} \tag{20}$$

Substitution of Eq. (20) in Eq. (19) results in the following form of the Hamiltonian:

$$H = \tfrac{1}{2}[x^T Q x - \boldsymbol{\lambda}^T B R^{-1} B^T \boldsymbol{\lambda}] + \boldsymbol{\lambda}^T[Ax] \tag{21}$$

The transversality condition is written as

$$\boldsymbol{\lambda}(t_f) = \frac{\partial J^*[x(t_f), t_f]}{\partial x(t_f)} = C^T\boldsymbol{\nu} \tag{22}$$

The next step for obtaining a feedback control law is an expansion of the cost-to-go. Ideally, it is desirable to express the cost function in terms of the current and final states. However, this is not possible directly because an expansion of this type does not lead to explicit terminal boundary conditions for the gains. This situation can be identified with the lack of identity transformations for certain types of generating functions in the canonical transformation approach.[23] Because the Hamiltonian is quadratic in x and $\boldsymbol{\lambda}$ and Eq. (22) suggests a linear dependence of $\boldsymbol{\lambda}$ on $\boldsymbol{\nu}$, the following quadratic form is written for $J^*[x(t), t]$:

$$J^*[x(t), t] = \tfrac{1}{2}x(t)^T S(t)x(t) + \boldsymbol{\nu}^T K(t)x(t) + \tfrac{1}{2}\boldsymbol{\nu}^T P(t)\boldsymbol{\nu} - \boldsymbol{\nu}^T\psi_f \tag{23}$$

Equation (23) is a modified version of a similar form given in Bryson.[26] The addition of the last term in Eq. (23) is motivated by its presence in Eq. (3). The matrices S, K, and P in the preceding equations are time-varying gain matrices of appropriate dimensions. Differential equations and the appropriate boundary conditions for the gains have to be determined. Note that S and P can be assumed to be symmetric. The following results are obtained by utilizing Eqs. (11) and (14) for the LQ problem:

$$\boldsymbol{\lambda} = \frac{\partial J^*}{\partial x} = Sx + K^T\boldsymbol{\nu} \tag{24}$$

$$\frac{\partial J^*}{\partial \boldsymbol{\nu}} = 0 = Kx + P\boldsymbol{\nu} - \psi_f \tag{25}$$

The boundary conditions for the preceding gains can be obtained by inspecting Eqs. (18) and (22), as given next:

$$S(t_f) = 0 \tag{26}$$

$$K(t_f) = C \tag{27}$$

$$P(t_f) = 0 \tag{28}$$

Note also that these boundary conditions result in the following intuitive result for the cost at the final time:

$$J^*[x(t_f), t_f] = 0 \tag{29}$$

The HJB equation can be written as

$$\frac{\partial J^*}{\partial t} = -\frac{1}{2}\left[x^T Q x - \frac{\partial J^*}{\partial x}^T BR^{-1}B^T \frac{\partial J^*}{\partial x}\right] - \frac{1}{2}\frac{\partial J^*}{\partial x}^T[Ax] - \frac{1}{2}x^T A^T \frac{\partial J^*}{\partial x} \tag{30}$$

The differential equations for the gains are obtained by utilizing Eq. (23) in Eq. (30) and collecting the coefficients of the linear and quadratic terms involving x and $\boldsymbol{\nu}$. These equations are

$$\dot{S} = -Q - SA - A^T S + SBR^{-1}B^T S \tag{31}$$

$$\dot{K} = -K(A - BR^{-1}B^T S) \tag{32}$$

$$\dot{P} = KBR^{-1}B^T K^T \tag{33}$$

The optimal control law can be obtained from Eq. (20), via Eq. (24). However, in order to implement the control, the constant $\boldsymbol{\nu}$ must be calculated from Eq. (25) at any time, other than the final time, as follows:

$$\boldsymbol{\nu} = -P(t)^{-1}[K(t)x(t) - \psi_f] \tag{34}$$

Sufficient conditions for a minimum, based on the existence of solutions to Eqs. (31–33), can be found in Bryson and Ho.[25] Significance of the normality condition $P(t) < 0$ for $t_0 \leq t \leq t_f$ can be seen from the evaluation of the value function by replacing ψ_f in Eq. (23) by using Eq. (25). The result is

$$J^*[x(t), t] = \tfrac{1}{2}x(t)^T S(t)x(t) - \tfrac{1}{2}\boldsymbol{\nu}^T P(t)\boldsymbol{\nu} \tag{35}$$

Because S is required to be the positive-definite solution to the Riccati equation, the normality condition ensures that the value function remains positive definite, as it should be.

For an open-loop solution to the problem, the initial costate is all that is required. The initial costate can be computed with the knowledge of the initial values of the gains and $\boldsymbol{\nu}$. In the absence of numerical errors, $\boldsymbol{\nu}$ is a constant, irrespective of the time t, at which it is computed by using Eq. (34). The optimal control problem can be solved by forward integration of the state and costate differential equations, once the initial costate vector is determined from Eq. (24). However, a feedback control implementation requires that the gains be stored and the control be implemented using the current state and time. A singularity exists in the feedback control implementation at the final time, where $\boldsymbol{\nu}$ cannot be calculated from Eq. (34). A practical fix to this problem is to stop the updating of $\boldsymbol{\nu}$ during the "blind time."[26]

The dynamic-programming approach for the derivation of the feedback control law is more direct than that of the sweep method, using the state and costate differential equations.[25] Furthermore, the computational burden for determining the optimal control is reduced (considerably more so for nonlinear problems because of the exploitation of symmetry of the gain tensors) by implementing the HJB methodology. The method just presented is especially attractive when the dimension of $\boldsymbol{\nu}$ is much smaller than that of x because the number of gains is reduced. In the next section, the procedure just described for the LQ problem is extended to the more general case of a nonlinear system with a nonlinear terminal constraint.

Series Solution Methodology for the Nonlinear Problem

Series solution methods are attractive for weakly nonlinear systems that are analytic. Convergence of the series solution is not guaranteed for highly nonlinear systems. Herein, it is assumed that the states are nondimensionalized such that the series solution is valid. Alternatively, the nonlinear perturbation dynamics close to the nominal optimal trajectory can be considered, in order to justify the use of the series solution. Required assumptions regarding smoothness of the value function and existence of solutions are also made.

Because the Hamiltonian is a function involving the states and costates, the cost-to-go function is expanded in a polynomial series involving x and $\boldsymbol{\nu}$ as shown here:

$$\begin{aligned}J^*[x(t), t] &= S_{1_{ij}}(t)x_i x_j + S_{2_{ijk}}(t)x_i x_j x_k + \cdots + K_{1_{pj}}(t)\nu_p x_j \\ &\quad + K_{2_{pqj}}(t)\nu_p\nu_q x_j + K_{3_{pij}}(t)\nu_p x_i x_j + \cdots - (\psi_f)_p\nu_p \\ &\quad + C_{1_{ipq}}(t)\nu_p\nu_q + C_{2_{ipqr}}(t)\nu_p\nu_q\nu_r + \cdots h.o.t \\ &\qquad i, j, k, l \ldots\ldots etc. = 1, 2, 3, \ldots, n \\ &\qquad p, q, r \ldots\ldots etc = 1, 2, 3, \ldots, p \leq n\end{aligned} \tag{36}$$

where $S_i(t), R_i(t), K_i(t), C_i(t) \ldots, i = 1, 2, 3, \ldots,$ are gain tensors (expressed using indicial notation), having time-dependent elements. This polynomial expansion of the cost-to-go is at the heart of the method presented in this paper.

Differential equations for the gain elements can be obtained by substituting Eq. (36) into the HJB equation [Eq. (9)] and collecting the coefficients of like powered terms involving x and $\boldsymbol{\nu}$. For example, it can be seen that $S_1(t)$ satisfies the familiar Riccati equation. Terminal boundary conditions on the gain variables can be postulated by using Eqs. (8) and (15). Once the gain elements are obtained via backward integration, the costate vector can be determined from Eq. (11), and $\boldsymbol{\nu}$ can be determined from Eq. (14) by using vector reversion of series.[28]

Generation of the equations for the gain elements is a tedious process, but it can be simplified by the use of symbolic manipulation programs like Mathematica® or Maple. Structured approaches to the series solution methodology have been presented in Refs. 6 and 7. It is convenient to combine x and $\boldsymbol{\nu}$ into a single extended vector for generating the gain equations. The equations required in this work were generated by using the software package, Maple.

Example with a Closed-Form Solution

The following, one-dimensional, example is considered to illustrate the application of the methodology:

Minimize:

$$J = \frac{1}{2}\int_0^{t_f} u^2\,\mathrm{d}t$$

Subject to:

$$\dot{x} = u, \qquad \psi = x(t_f) + \varepsilon x(t_f)^2 = \psi_f$$
$$x(0), \psi_f, \qquad \text{and} \qquad t_f = \text{given} \tag{37}$$

where ε is a small parameter. It is assumed that the constraint is regular, that is, at least one real solution exists for $x(t_f)$. This problem is quite simple if the constraint is solved for $x(t_f)$ a priori. Without an explicit solution for the final state, assuming that two real roots exist, it will be desirable for the algorithm to choose the final condition that involves the least cost.

The HJB equation for the example problem is

$$\frac{\partial J^*[x(t), t]}{\partial t} = \frac{1}{2}J^{*2}_x \tag{38}$$

subject to

$$J^*_x[x(t_f), t_f] = \boldsymbol{\lambda}(t_f) = \boldsymbol{\nu}\psi_x[x(t_f)] = \boldsymbol{\nu}[1 + 2\varepsilon x(t_f)] \tag{39}$$

The cost-to-go is expanded using a power series as follows:

$$\begin{aligned}J^*[x(t), t] &= S_1x^2 + S_2x^3 + \cdots + K_1\boldsymbol{\nu}x + K_2\boldsymbol{\nu}^2x + K_3\boldsymbol{\nu}x^2 \\ &\quad + \cdots - \psi_f\boldsymbol{\nu} + C_1\boldsymbol{\nu}^2 + C_2\boldsymbol{\nu}^3 + \cdots\end{aligned} \tag{40}$$

Thus the expansions for the costate and the constraint are

$$\boldsymbol{\lambda} = 2S_1x + 3S_2x^2 + \cdots + K_1\boldsymbol{\nu} + K_2\boldsymbol{\nu}^2 + 2K_3\boldsymbol{\nu}x + \cdots \tag{41}$$

and

$$K_1 x + 2K_2 \boldsymbol{\nu} x + K_3 x^2 + \cdots + 2C_1 \boldsymbol{\nu} + 3C_2 \boldsymbol{\nu}^2 + \cdots = \psi_f \quad (42)$$

Equation (42) can be used to solve for $\boldsymbol{\nu}$, once the gains have been computed. The following system of differential equations for the gains is obtained by substituting Eq. (40) into Eq. (38) and utilizing Eq. (39):

$$\dot{S}_1 = 2S_1^2, \qquad S_1(t_f) = 0, \qquad \dot{S}_2 = 6S_1S_2, \qquad S_2(t_f) = 0$$

$$\cdot\cdot$$

$$\dot{K}_1 = 2S_1K_1, \qquad K_1(t_f) = 1, \qquad \dot{K}_2 = 2S_1K_2 + 2K_1K_3$$

$$K_2(t_f) = 0, \qquad \dot{K}_3 = 3S_2K_1 + 4S_1K_3, \qquad K_3(t_f) = \varepsilon$$

$$\cdot\cdot\cdot\cdot$$

$$\dot{C}_1 = \tfrac{1}{2}K_1^2, \qquad C_1(t_f) = 0, \qquad \dot{C}_2 = K_1K_2, \qquad C_2(t_f) = 0$$

$$\cdot\cdot\cdot\cdot\cdot \quad (43)$$

The preceding differential equations can be solved analytically, and the following is the series representation for the control obtained from Eq. (41):

$$\boldsymbol{\lambda} = -u = \left[\boldsymbol{\nu} + 2\varepsilon(t - t_f)\boldsymbol{\nu}^2 + 4\varepsilon^2(t - t_f)^2\boldsymbol{\nu}^3 + 8\varepsilon^3(t - t_f)^3\boldsymbol{\nu}^4 \ldots\right]\psi_x[x(t)] \quad (44)$$

The preceding series is a binomial expansion, and it is recognized that as more terms in the series are taken the resulting control law takes the form:

$$u = -\boldsymbol{\nu}\left[\frac{1}{1 - 2\varepsilon(t - t_f)\boldsymbol{\nu}}\right]\psi_x = -\boldsymbol{\nu}\left[\frac{1}{1 - 2\varepsilon(t - t_f)\boldsymbol{\nu}}\right](1 + 2\varepsilon x) \quad (45)$$

Hence, the series in Eq. (44) converges if $|2\varepsilon(t - t_f)\boldsymbol{\nu}| < 1$. Equation (42) can be reduced to the form shown next by using the solutions to Eq. (43):

$$\left[\boldsymbol{\nu}(t - t_f) + 3\varepsilon(t - t_f)^2\boldsymbol{\nu}^2 + 8\varepsilon^2(t - t_f)^3\boldsymbol{\nu}^3 + 20\varepsilon^3(t - t_f)^4\boldsymbol{\nu}^4 \ldots\right] = -\left\{\frac{\psi[x(t)] - \psi_f}{\psi_x^2[x(t)]}\right\} \quad (46)$$

which can also be written compactly as

$$\boldsymbol{\nu}(t - t_f)\left\{\frac{1 - \varepsilon(t - t_f)\boldsymbol{\nu}}{[1 - 2\varepsilon(t - t_f)\boldsymbol{\nu}]^2}\right\} = -\left\{\frac{\psi[x(t)] - \psi_f}{\psi_x^2[x(t)]}\right\} \quad (47)$$

On further inspection, the preceding result simplifies to

$$[1 - 2\varepsilon(t - t_f)\boldsymbol{\nu}]^2 = \frac{(1 + 2\varepsilon x)^2}{[1 + 4\varepsilon\psi_f]} \qquad \varepsilon \neq 0$$

$$\boldsymbol{\nu} = -\left(\frac{x - \psi_f}{t - t_f}\right), \qquad \varepsilon = 0 \quad (48)$$

Hence, $\boldsymbol{\nu}$ can be computed from Eq. (46), by using series reversion or directly, from Eq. (48). The process of series reversion yields an unique solution to $\boldsymbol{\nu}$:

$$\boldsymbol{\nu} = F - 3\varepsilon(t - t_f)F^2 + 10\varepsilon^2(t - t_f)^2F^3 - 35\varepsilon^3(t - t_f)^3F^4 + 126\varepsilon^4(t - t_f)^4F^5 - 462\varepsilon^5(t - t_f)^5F^6 + \cdots\cdot \quad (49)$$

where

$$F = -\frac{1}{(t - t_f)}\left\{\frac{\psi[x(t)] - \psi_f}{\psi_x^2[x(t)]}\right\}$$

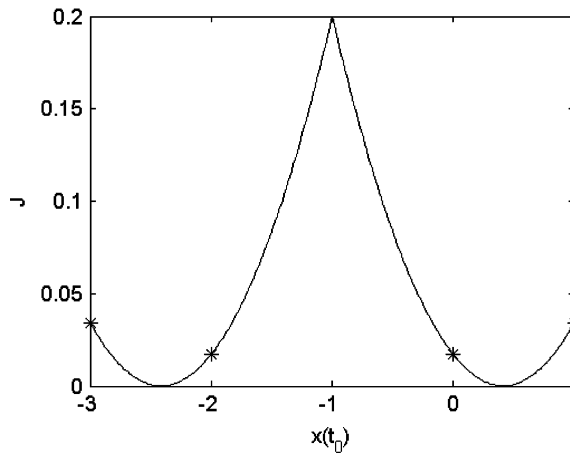

Fig. 1 Optimal value function: comparison of the results of the series and exact solutions.

On the other hand, the proper solution to Eq. (48) must be selected to obtain the optimal value function. Theoretically, $\boldsymbol{\nu}$ is a constant, but its evaluation from Eq. (46) shows variation with respect to time because of truncation of the series. Note that if $\psi_x = 0$ initially, then $\boldsymbol{\nu}$ cannot be computed from Eq. (49). This special case corresponds to the existence of a point from which two equally optimal solutions exist (Darboux point[29]). The series solution, in the form just presented, is not valid for this special initial condition because the requirement for convergence of the series is violated. When posed as a fixed final-state problem, the difficulty with the special initial condition just mentioned disappears. However, one must solve the problem for multiple possible final states and choose the minimum-cost solution (not a practical method of solution for more difficult problems). The control gain for either method approaches infinity, as the final time is reached. The advantage of the series solution method is that it automatically chooses the optimal control, in the neighborhood of the terminal constraint, when the convergence criterion is satisfied and the required number of terms is included in the series expansion.

For completeness, the exact, value function for the preceding problem is given next:

$$J^*[x(t), t] = \boldsymbol{\nu}\left[\frac{\psi + (\nu/2)(t - t_f)}{1 - 2\varepsilon(t - t_f)} - \psi_f\right] \quad (50)$$

where $\boldsymbol{\nu}$ can be evaluated from Eq. (48). This result is used to validate the series solution method.

Consider an example with $\varepsilon = 0.5, t_f = 5$ s, and $\psi_f = 0.5$. Figure 1 shows the exact, bimodal cost function for this problem, evaluated at the initial time, $t = t_0 = 0$. Results of the series solution procedure for four different initial conditions are superimposed on this figure. These results were obtained by integrating the closed-loop system with the feedback control law given by Eq. (44), with $\boldsymbol{\nu}$ evaluated from Eq. (49). Six terms were used in the expansions in each of the preceding equations. As shown in Fig. 1, values of the cost function are reproduced accurately at the selected points. The six-term series solution does not converge when the initial state is in the range (−2,0).

Other Examples

The following, two-dimensional, example is presented to further illustrate the application of the series solution methodology. Unlike for the preceding example with an analytical solution for the gain differential equations, the gain equations for the examples presented in this section were integrated backwards in time to obtain their initial values. In the actual implementation of the control law, the gain equations were integrated forward in time, along with the state equations, using the known initial conditions. Although the computational burden increases with such an implementation, storage of the time-varying gains is avoided. A different implementation is utilized in the next section. In general, especially when integrating over large time intervals, some form of a symplectic integrator should be utilized, as suggested in Ref. 22. In all of the examples presented next, the value function was expanded to fourth order, and the series reversion process to determine $\boldsymbol{\nu}$ was also implemented to fourth order. Furthermore, because $\boldsymbol{\nu}$ cannot be determined at the final time, the simulation was stopped just prior to reaching the final time. For each example, the corresponding open-loop optimal solution, obtained by using a shooting method, is also presented.

Equations (51), (52), and (53), respectively, show the dynamical model, constraint, and performance index:

$$\ddot{x} = -\left(m_1 x + m_2 x^2 + m_3 x^3\right) + Bu \tag{51}$$

$$\psi = \left[\varepsilon_1 x(t_f) + \varepsilon_2 \dot{x}(t_f) + \varepsilon_3 x(t_f)^2 + \varepsilon_4 \dot{x}(t_f)^2 + \varepsilon_5 x(t_f)\dot{x}(t_f)\right] = \psi_f \tag{52}$$

$$J = \int_{t_0}^{t_f} \left(Q_{11}x^2 + Q_{22}\dot{x}^2 + Ru^2\right) \mathrm{d}t \tag{53}$$

Two cases are considered next by choosing different sets of parameters for the model.

Case A: Highly Nonlinear System and Nonlinear Terminal Constraint

This example is parameterized by the following choices for the constants in Eqs. (51–53):

$$m_1 = 1, \quad m_2 = 0.8, \quad m_3 = 0.75, \quad \varepsilon_1 = 0.9$$

$$\varepsilon_2 = 0.6, \quad \varepsilon_3 = 0.6, \quad \varepsilon_4 = 0.3, \quad \varepsilon_5 = 0.1$$

$$Q_{11} = Q_{22} = B = R = 1, \quad t_0 = 0, \quad t_f = 3$$

$$x(t_o) = [0.1; 0.2], \quad \psi_f = 0.5$$

Figure 2a shows the open-loop and feedback control histories for the nominal and perturbed initial conditions. The feedback control histories track their open-loop counterparts closely. Deviations between the open-loop and feedback control histories are evident because of the approximations introduced by the series expansion and the series reversion process. The results for the perturbed initial conditions were obtained by utilizing forward integration only, because the initial conditions of the gains were known from the computations for the nominal problem. The error in satisfying the terminal constraint is approximately 4.8×10^{-5}, for a 10% change in the initial state from its nominal value. The simulation is stopped slightly before reaching the final time of 3 s because of the singularity at the final time. Figure 2b shows the open-loop and feedback phase portraits for various initial conditions. Even though there is some deviation in the midcourse, the terminal constraint is satisfied quite accurately. Figure 2c shows the convergence of the first state toward its optimal solution, as more terms in the series are included. Third-order series solution for the control is seen to be much better than the lower-order control approximations.

Another interesting observation is the existence of multiple open-loop solutions obtained for the same initial conditions, depending on the initial guesses for the costates. Because the shooting method utilizes necessary conditions only, the result cannot be guaranteed to be a local minimum unless tested further using second-order conditions. Figure 2d shows three trajectories from the same initial

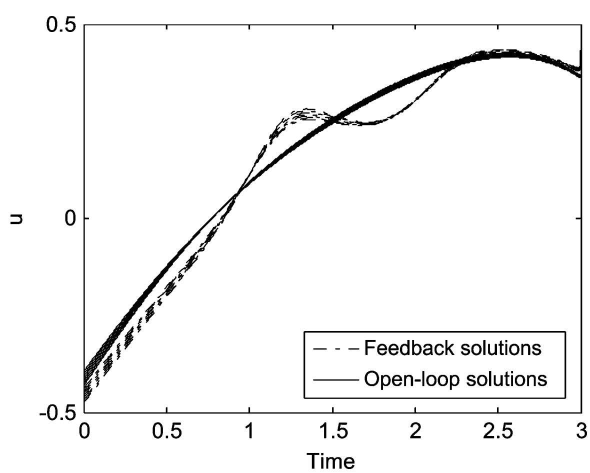

Fig. 2a Case A: open-loop and feedback control histories.

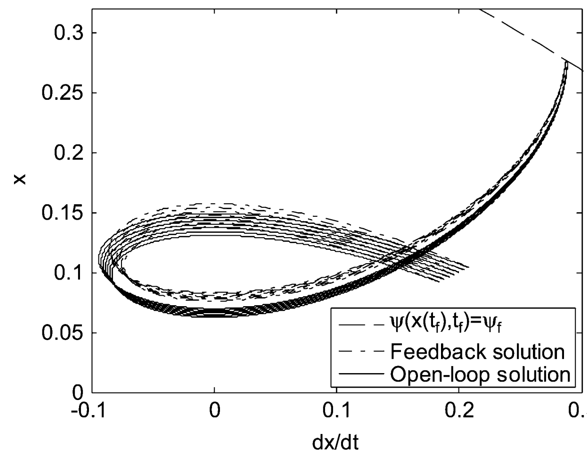

Fig. 2b Case A: phase portraits with various I.C. (Nominal initial conditions (I.C.) is perturbed within the range [−10 10] %.)

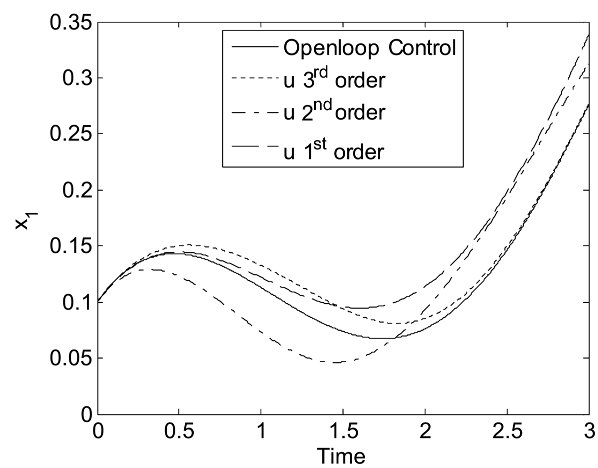

Fig. 2c Case A: convergence of the state trajectory as a function of the order of approximation.

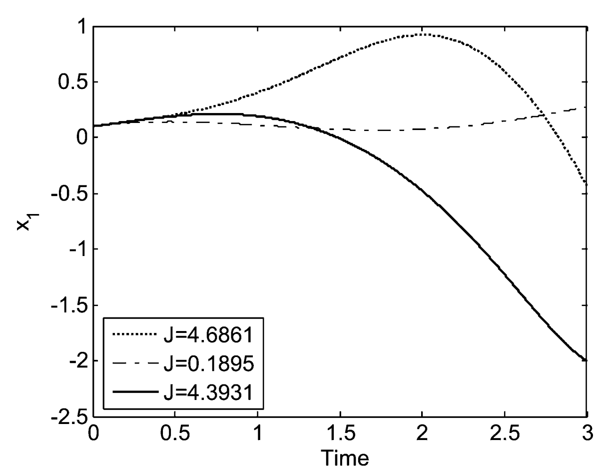

Fig. 2d Multiple open-loop extremals for the same state I.C. but with different initial costate guesses.

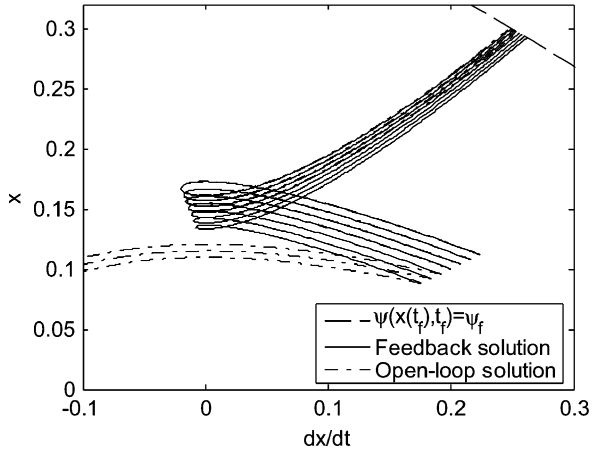

Fig. 3 Case B: phase portraits for the feedback and open-loop solutions for different state I.C.

conditions for the states, each with a different cost function. However, each feedback-controlled trajectory, for all of the initial conditions, is associated with the lowest cost solution, for the respective case.

Case B: Unstable Nonlinear Plant with a Nonlinear Terminal Constraint

Parameters for this example are given here:

$$m_1 = -1, \qquad m_2 = 0.8, \qquad m_3 = 0.75, \qquad \varepsilon_1 = 0.9$$

$$\varepsilon_2 = 0.6, \qquad \varepsilon_3 = 0.6, \qquad \varepsilon_4 = 0.3, \qquad \varepsilon_5 = 0.1$$

$$Q_{11} = Q_{22} = B = R = 1, \qquad t_0 = 0, \qquad t_f = 3$$

$$x(t_o) = [0.1; 0.2], \qquad \psi_f = 0.5$$

The linear part of this system is unstable for this example. Figure 3 shows the phase portraits of the optimal trajectories, both closed loop and open loop. As can be seen, all of the feedback and some of the open-loop trajectories, for neighboring initial conditions, lie close to each other. Some of the open-loop trajectories converge to higher-cost solutions, depending on the initial guesses for the costate values. As mentioned before, the same initial guess was used to obtain the solutions for all of the initial conditions.

Application to Optimal Detumbling of a Spacecraft

The series solution methodology is finally applied to the problem of optimal detumbling maneuvers of a rigid asymmetric spacecraft. The Euler equations for a rigid body are

$$\dot{\omega} = I^{-1}[\tilde{\omega} I \omega + u] \tag{54}$$

where $\omega = (\omega_1, \omega_2, \omega_3) \in \Re^{3\times 1}$ and $u = (u_1, u_2, u_3) \in \Re^{3\times 1}$ are the angular velocity and input control torque vectors, respectively. The moment of inertia matrix $I \in \Re^{3\times 3}$ is represented in the principal axes system as

$$I = \begin{bmatrix} I_1 & 0 & 0 \\ 0 & I_2 & 0 \\ 0 & 0 & I_3 \end{bmatrix} \tag{55}$$

The angular velocity cross-product operator is represented by $\tilde{\omega}$. The principal moments of inertia for this example are given here[30]:

$$I_1 = 86.24\,\text{kg-m}^2, \qquad I_2 = 85.07\,\text{kg-m}^2, \qquad I_3 = 113.59\,\text{kg-m}^2$$

The performance index is selected to be of the form

$$J = \frac{1}{2}\int_0^T \left(\omega^T Q \omega + u_1^2 + u_2^2 + u_3^2\right) \mathrm{d}t \tag{56}$$

where Q is a positive-definite weight matrix.

The solution presented was developed based upon an expansion of the cost-to-go to fourth-order and a second-order solution for the series reversion.[28] Series reversion to higher order was not carried out because of excessive computational burden. Series reversion to second order was deemed sufficient for this problem, after inspecting the results. All of the time-dependent gains were stored at discrete points of time during backward integration, and the stored gains were utilized to calculate the feedback control and the trajectory, during the forward integration process. A cubic-spline-interpolation technique available in MATLAB® was used to determine gains at intermediate time points between two stored data points. Results obtained for a specific numerical example are presented next.

In the example considered, the final time was selected to be 2 s, and the weight matrix in the performance index is selected as the moment of inertia matrix, that is, $Q = I$. The initial angular velocities selected are as shown next:

$$\omega_1(0) = -0.4\ \text{rad/s}, \qquad \omega_2(0) = 0.8\ \text{rad/s}, \qquad \omega_3(0) = 2\ \text{rad/s}$$

This example is not very realistic because of the high torque requirements but is used to demonstrate the accuracy of our methodology. Figure 4a shows the trajectories obtained by using the feedback solution, and for comparison the open-loop solutions are also presented on the same graph. Again, there is a very close match between the two solutions. Figure 4b shows the control torques, and slight deviations between the respective closed-loop and open-loop profiles are noticed. These deviations can be attributed to the very high initial angular velocities and the short maneuver time. Even though the differences between the respective control histories are noticeable, the open-loop and feedback controlled state trajectories are almost identical. As an additional check, the Lagrange multiplier $\boldsymbol{\nu}$, computed along the trajectory, is plotted in Fig. 4c. Ideally, this vector

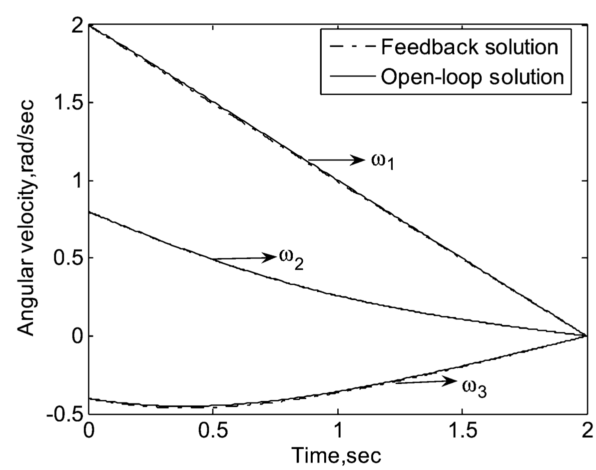

Fig. 4a Angular velocities for the spin maneuver.

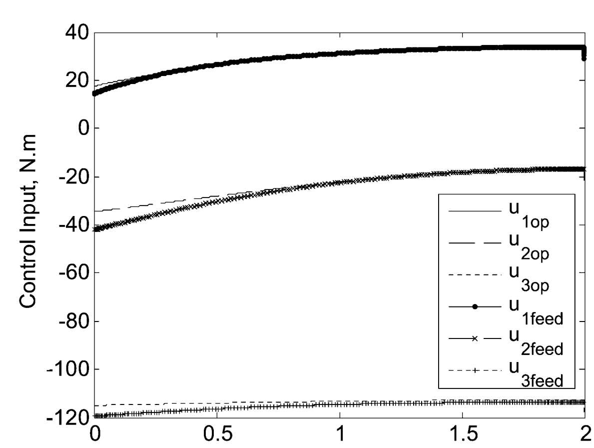

Fig. 4b Feedback and open-loop control inputs.

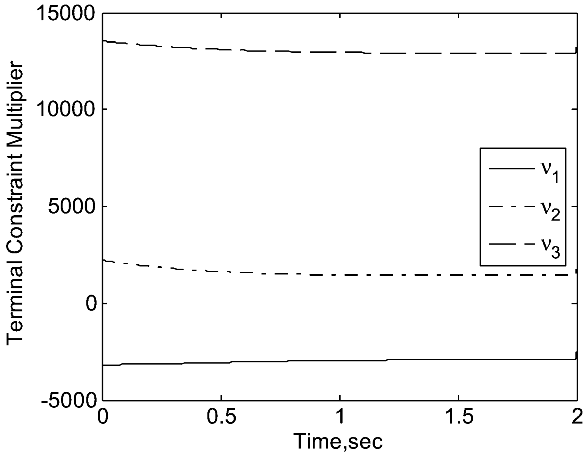

Fig. 4c Lagrange multipliers associated with the terminal constraints.

should be constant, but it is not so because of series truncation. As mentioned earlier, very near the final time, the Lagrange multipliers exhibit sharp changes. Hence the forward integration process was stopped just before reaching the final time.

Conclusions

This paper presents a series solution method for determining finite-horizon optimal feedback control laws for nonlinear systems with nonlinear terminal constraints. Expansion of the value function as a polynomial series, in terms of the states and the constraint Lagrange multipliers, is utilized for the solution of the Hamilton–Jacobi–Bellman equation. The method is applicable to analytic systems and constraints. This method is more direct and accomplished in a single step rather than in two steps as in the sweep method. The resulting optimal control law performs exceedingly well, as demonstrated on some numerical examples, including the optimal spacecraft detumbling problem. Many practical guidance problems require the solutions to fixed final time, terminally constrained nonlinear optimal control problems. There might be significant advantages to neighboring optimal control methods based on nonlinear realizations of the perturbation dynamics.

Acknowledgments

The authors wish to thank James D. Turner for his help with the implementation of the vector series reversion algorithm.

References

[1]Debs, A. S., and Athans, M., "On the Optimal Angular Velocity Control of Asymmetrical Space Vehicles," *IEEE Transactions on Automatic Control*, Vol. 14, No. 1, Feb. 1969, pp. 80–83.

[2]Al'brekht, E. G., "On the Optimal Stabilization of Nonlinear Systems," *Journal of Applied Mathematics and Mechanics*, Vol. 25, No. 5, 1962, pp. 1254–1266.

[3]Lukes, D. L., "Optimal Regulation of Nonlinear Dynamical Systems," *SIAM Journal of Control*, Vol. 7, No. 1, 1969, pp. 75–100.

[4]Dabbous, T. E., and Ahmed, N. U., "Nonlinear Optimal Feedback Regulation of Satellite Angular Momenta," *IEEE Transactions on Aerospace and Electronic Systems*, Vol. AES-18, No. 1, 1982, pp. 2–10.

[5]Carrington, C. K., and Junkins, J. L., "Optimal Nonlinear Feedback Control for Spacecraft Attitude Maneuvers," *Journal of Guidance, Control, and Dynamics*, Vol. 9, No. 1, 1986, pp. 99–107.

[6]Yoshida, T., and Laparo, K. A., "Quadratic Regulatory Theory for Analytic Nonlinear Systems with Additive Controls," *Automatica*, Vol. 25, No. 4, 1989, pp. 531–544.

[7]Huang, J., and Lin, C. F., "Numerical Approach to Computing Nonlinear H-Infinity Control Laws," *Journal of Guidance, Control, and Dynamics*, Vol. 18, No. 5, 1995, pp. 989–994.

[8]Tsiotras, P., Corless, M., and Rotea, M., "An L2 Disturbance Attenuation Solution to the Nonlinear Benchmark Problem," *International Journal of Robust and Nonlinear Control*, Vol. 8, 1998, pp. 311–330.

[9]Garrard, W. L., McClamroch, N. H., and Clark, L. G., "An Approach to Suboptimal Feedback Control of Nonlinear Systems," *International Journal of Control*, Vol. 5, No. 5, 1967, pp. 425–435.

[10]Cloutier, J. R., "State-Dependent Riccati Equation Techniques," *Proceedings of the American Control Conference*, American Automatic Control Council, Dayton, OH, June 1997, pp. 932–936.

[11]Tewari, A., "Optimal Nonlinear Spacecraft Attitude Control Through Hamilton-Jacobi Formulation," *Journal of the Astronautical Sciences*, Vol. 50, No. 1, 2002, pp. 99–112.

[12]Sharma, R., and Tewari, A., "Optimal Nonlinear Tracking of Spacecraft Attitude Maneuvers," *IEEE Transactions on Control Systems Technology*, Vol. 12, No. 5, 2004, pp. 677–682.

[13]Cimen, T., and Banks, S. P., "Global Optimal Feedback Control for General Nonlinear Systems with Nonquadratic Performance Criteria," *Systems and Control Letters*, Vol. 53, No. 5, 2004, pp. 327–424.

[14]Xin, M., Balakrishnan, S. N., Stansbery, D. T., and Ohlmeyer, E. J., "Nonlinear Missile Autopilot Design with $\theta - D$ Technique," *Journal of Guidance, Control, and Dynamics,* Vol. 27, No. 3, 2004, pp. 406–417.

[15]Vadali, S. R., and Junkins, J. L., "Optimal Open Loop and Stable Feedback Control of Rigid Spacecraft Attitude Maneuvers," *The Journal of Astronautical Sciences*, Vol. 32, No. 2, 1984, pp. 105–122.

[16]Kristic, M., and Tsiotras, P., "Inverse Optimal Stabilization of a Rigid Spacecraft," *IEEE Transactions on Automatic Control Systems*, Vol. 44, No. 5, 1999, pp. 1042–1048.

[17]Beard, R., Saridis, G., and Wen, J., "Galerkin Approximation of the Generalized Hamilton-Jacobi-Bellman Equation," *Automatica*, Vol. 33, No. 12, 1997, pp. 2159–2177.

[18]Beard, R., and McLain, T., "Successive Galerkin Approximation Techniques for Nonlinear Optimal and Robust Control," *International Journal of Control*, Vol. 71, No. 5, 1998, pp. 717–748.

[19]Curtis, J. W., and Beard, R. W., "Successive Collocation: An Approximation to Optimal Nonlinear Control," *Proceedings of the American Control Conference*, American Automatic Control Council, Dayton, OH, June 2001, pp. 3481–3485.

[20]Park, C., and Tsiotras, P., "Sub-Optimal Feedback Control Using a Successive Wavelet-Galerkin Algorithm," *Proceedings of the American Control Conference*, American Automatic Control Council, Dayton, OH, June 2003, pp. 1926–1931.

[21]Park, C., and Tsiotras, P., "Approximations to Optimal Feedback Control Using a Successive Wavelet Collocation Algorithm," *Proceedings of the American Control Conference*, American Automatic Control Council, Dayton, OH, June 2003, pp. 1950–1955.

[22]Guibout, V. M., and Scheeres, D. J., "Solving Relative Two-Point Boundary Value Problems: Spacecraft Formation Flight Transfer Applications," *Journal of Guidance, Control, and Dynamics*, Vol. 27, No. 4, 2004, pp. 693–704.

[23]Park, C., and Scheeres, D. J., "Solutions of Optimal Feedback Control Problems with General Boundary Conditions Using Hamiltonian Dynamics and Generating Functions," *Proceedings of the American Control Conference*, American Automatic Control Council, Dayton, OH, June 2004, pp. 679–684.

[24]Park, C., and Scheeres, D. J., "Extended Applications of Generating Functions to Optimal Feedback Control Problems," *Proceedings of the American Control Conference*, American Automatic Control Council, Dayton, OH, June 2005, pp. 852–857.

[25]Bryson, A. E., and Ho, Y.-C., *Applied Optimal Control*, Hemisphere Washington, DC, 1975, pp. 217–236.

[26]Bryson, A. E., *Dynamic Optimization*, Addison Wesley Longman, Menlo Park, CA, 1999, pp. 262–325.

[27]Dreyfus, S., *Dynamic Programming and the Calculus of Variations*, Academic Press, New York, 1965.

[28]Turner, J. D., "Generalized Gradient Search and Newton's Method for Multilinear Algebra Root Solving and Optimization Applications," *Proceedings of John L. Junkins Astrodynamics Symposium, Advances in the Astronautical Sciences*, Vol. 115, Univelt Publishers, San Diego, CA, 2003, pp. 55–78.

[29]Mereau, P. M., and Powers, W. F., "The Darboux Point," *Journal of Optimization Theory and Applications*, Vol. 17, No. 5/6, 1975, pp. 545–559.

[30]Junkins, J. L., and Turner, J. D., *Optimal Spacecraft Rotational Maneuvers*, Elsevier, New York, 1986, pp. 243, 244.