JOURNAL OF GUIDANCE, CONTROL, AND DYNAMICS
Vol. 17, No. 5, September–October 1994

Hybrid Approach to Solution of Optimal Control Problems

Anthony J. Calise* and Martin S. K. Leung†
Georgia Institute of Technology, Atlanta, Georgia 30332

This paper proposes a solution approach for nonlinear optimization problems that seeks to combine the desirable features of analytic methods that are based on use of simplified models and numerical methods that use piecewise elementary interpolation functions to represent the solution. The approach is illustrated by a combination of regular perturbation analysis and the method of collocation, although other combinations of model simplification and numerical solution methods can also be envisioned. A simple fourth-order nonlinear system is used to illustrate the conceptual approach for several possible levels of approximation.

Introduction

THE method of regular perturbation analysis has been proposed as an approach to development of real-time guidance algorithms.[1–5] This approach allows approximation of the optimal control solution by expanding the solution in an asymptotic power series in a small parameter either appearing naturally in the differential equations of motion or artificially introduced (through engineering judgment) and used as a bookkeeping parameter for the analysis. The zero-order problem is a simplified optimal control problem, which in some practical applications can be solved analytically. Crucial to the success of the method is that the optimal solution is reasonably approximated by the zero-order solution, so that the addition of first- or higher order corrections to the series solution (which usually is not convergent) results in an improvement in accuracy. The higher order corrections involve the solution of nonhomogeneous linear time-varying differential equations that may be solved by quadrature. The approach has great success when applied to systems with small nonlinear terms, so that the zero-order problem is linear.[1,3] Also, in certain applications a state transition matrix may be determined for the first- and higher order corrections, further facilitating the solution process. The major limitation in guidance applications appears to be that significant nonlinearities, such as the aerodynamic effects, must be neglected in the zero-order problem in order to obtain an analytic solution,[4,5] which is also nonlinear even in the absence of aerodynamic effects. It turns out in this case that the zero-order problem is not sufficiently close to the original problem and the solution begins to diverge even when a first-order correction is attempted.[6] A second drawback inherent in any attempt at analysis by model simplification is that a significant amount of reanalysis is required when even a minor change in the optimal control problem formulation is made.

Collocation[7,8] is a general method for obtaining an approximate solution of differential equations. It involves choosing simple interpolating functions and enforcing the interpolatory constraints at specific points within segments of the time interval to evaluate the unknown coefficients. Thus, when applied to an optimal control problem, it reduces the associated two-point boundary-value problem to a set of coupled nonlinear algebraic equations. Collocation methods have the advantages that they are simple to use for a wide variety of optimization problems, and their accuracy can be improved by increasing the number of segments used in the approximation. The major disadvantages are that there is no general guarantee that the numerical methods employed will successfully solve the nonlinear programming problem under all circumstances, and the dimension of the problem increases proportionately with the number of segments.

It is apparent from the above discussion that the advantages of analytic and numerical methods are in many respects complementary in the sense that if the advantages can be combined in some way, then most of the important disadvantages (from the viewpoint of real-time applications) can be reduced. This paper proposes two of possibly many ways to obtain such a hybrid methodology, with the potential for use in the development of real-time optimal guidance algorithms. The first approach uses the method of regular expansion to improve upon a collocation solution, thereby reducing the error for a given number of segments. The second approach improves upon the first by using both regular expansion and analytic methods to identify more intelligent interpolating functions in the collocation method, again with the objective of improving the level of accuracy without increasing the number of segments. An application of the second approach to launch vehicle guidance is treated in a companion paper to this one, and further details regarding this application may be found in Refs. 6 and 9.

We begin with the formulation of a regularly perturbed optimal control problem. This is followed by a discussion on the use of collocation for generating approximate solutions and how it can be used in the setting of regular perturbation analysis. The hybrid use of the two techniques is then illustrated in a Duffing equation example, in which several interpolating functions of increasing degrees of sophistication are used and evaluated. First- and second-order corrections in the asymptotic series approximation are also included to complete the example.

Regular Perturbation in Optimal Control

The optimal control problem formulation considered here is to minimize a performance index that is a function of the terminal states and time, subject to nonlinear dynamic constraints:

$$J = \min_u [\phi(x, t)]|_{t_f} \tag{1}$$

$$\dot{x} = f(x, u, t) + \epsilon g(x, u, t), \qquad x(t_0) = x_0, \qquad t \in [t_0, t_f] \tag{2}$$

and the terminal time constraints $\psi_i[x(t_f)] = 0, i = 1, \ldots, p \leq n$. In Eq. (2), x is an n-dimensional state vector and u is an m-dimensional control vector. The expansion parameter ϵ may be artificially inserted to signify the presence of small nonlinear effects. The Hamiltonian and transversality condition are given by

$$H = \lambda^T(f + \epsilon g), \qquad H(t_f) = -\Phi_t|_{t_f}, \qquad \Phi = \phi + \nu^T \psi \tag{3}$$

The costate equations and associated boundary conditions are

$$\dot{\lambda} = -H_x, \qquad \lambda(t_f) = \Phi_x|_{t_f} \tag{4}$$

Presented as Paper 92-4304 at the AIAA Guidance, Navigation, and Control Conference, Aug. 10–12, 1992; received March 9, 1992; revision received Sept. 27, 1993; accepted for publication Nov. 29, 1993.

*Professor, School of Aerospace Engineering. Fellow AIAA.
†Currently Project Manager, Business Development, Champion Technology Ltd., 9/F Kantone Centre, 1 Ning Foo Street, Chai Wan, Hong Kong. Member AIAA.

where the subscript is used to denote partial differentiation. In the absence of control constraints, the optimal control satisfies

$$H_u = \lambda^T(f_u + \epsilon g_u) = 0 \tag{5}$$

assuming that $H_{uu} > 0$. In the above final time is free; thus we introduce a new independent variable $\tau = (t - t_o)/T$, where $T = t_f - t_o$ and rewrite the necessary conditions of Eqs. (2–4) in the following equivalent form:

$$x' = H_\lambda T, \qquad x(\tau = 0) = x_0, \qquad \psi[x(\tau = 1)] = 0 \tag{6}$$

$$\lambda' = -H_x T, \qquad \lambda(\tau = 1) = \Phi_x|_{\tau=1} \tag{7}$$

$$T' = 0 \tag{8}$$

$$\begin{gathered} H = \lambda^T[f(x, u, \tau T + t_0) + \epsilon g(x, u, \tau T + t_0)] \\ H(\tau = 1) = -\Phi_t|_{\tau=1} \end{gathered} \tag{9}$$

where $(\bullet)'$ denotes $d(\bullet)/d\tau$. In a regular perturbation analysis, the objective is to approximate the solution to Eqs. (6–9) by an asymptotic series in x, λ, u, and T as follows:

$$\begin{aligned} x &= x_0 + \epsilon x_1 + \epsilon^2 x_2 + \cdots \\ \lambda &= x_0 + \epsilon\lambda_1 + \epsilon^2\lambda_2 + \cdots \\ u &= u_0 + \epsilon u_1 + \epsilon^2 u_2 + \cdots \\ T &= T_0 + \epsilon T_1 + \epsilon^2 T_2 + \cdots \end{aligned} \tag{10}$$

Assume the functions f, g, ϕ, ψ have piecewise continuous derivatives up to order at least $K + 1$, where K is the order of approximation. Using the Taylor series formula, a finite series approximation is constructed according to

$$F\left(\sigma_0 + \sum_{k=1}^{K} \sigma_k \epsilon^k\right) = F(\sigma_0) + \epsilon \left.\frac{\mathrm{d}F}{\mathrm{d}\sigma}\right|_{\sigma_0} \sigma_1 + \epsilon^2\left(\left.\frac{\mathrm{d}F}{\mathrm{d}\sigma}\right|_{\sigma_0} \sigma_2 + \frac{1}{2!}\left.\frac{\mathrm{d}^2F}{\mathrm{d}\sigma^2}\right|_{\sigma_0} \sigma_1^2\right) + \cdots \tag{11}$$

where $\sigma = \{x, \lambda, v, u, T\}$. Substituting the series representation for each of the variables in Eqs. (6–9) and equating like powers in ϵ, we obtain the zero- and higher order necessary conditions. To zero order we have

$$\begin{gathered} \frac{\partial x_0}{\partial \hat{t}} = \frac{\partial H_0}{\partial \lambda_0}, \qquad x_0(t_0) = x_0, \qquad \psi[x_0(t_0 + T_0)] = 0 \\ \frac{\partial \lambda_0}{\partial \hat{t}} = \frac{-\partial H_0}{\partial x_0}, \qquad \lambda_0(\hat{t} = t_0 + T_0) = \left.\frac{\Phi(x_0, \hat{t})}{\partial x_0}\right|_{\hat{t}=t_0+T_0} \\ \frac{\partial H_0}{\partial u_0} = 0 \\ H_0 = \lambda^T f(x_0, u_0, \hat{t}), \qquad H_0(\hat{t} = t_0 + T_0) = -\left.\frac{\Phi(x_0, \hat{t})}{\partial \hat{t}}\right|_{\hat{t}=t_0+T_0} \end{gathered} \tag{12}$$

In Eq. (12), the new independent variable $\hat{t} = \tau T_0 + t_0$ has been introduced, where it should be noted that in the zero-order problem $T = T_0$.

It can be shown that all of the higher order problems involve the solution of dynamic equations of the form

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}\hat{t}}\begin{bmatrix} x_k \\ \lambda_k \end{bmatrix} &= \begin{bmatrix} A_{11}(x_0, \lambda_0, T_0) & A_{12}(x_0, \lambda_0, T_0) \\ A_{21}(x_0, \lambda_0, T_0) & A_{22}(x_0, \lambda_0, T_0) \end{bmatrix}\begin{bmatrix} x_k \\ \lambda_k \end{bmatrix} \\ &+ \frac{T_k}{T_0}\begin{bmatrix} C_1(x_0, \lambda_0, T_0) \\ C_2(x_0, \lambda_0, T_0) \end{bmatrix} \\ &+ \begin{bmatrix} P_{1k}(x_0, \lambda_0, T_0, \ldots, x_{k-1}, \lambda_{k-1}, T_{k-1}) \\ P_{2k}(x_0, \lambda_0, T_0, \ldots, x_{k-1}, \lambda_{k-1}, T_{k-1}) \end{bmatrix} \end{aligned} \tag{13}$$

where

$$\begin{aligned} A_{11} &= f_x - f_u[(f_u^T\lambda)_u]^{-1}(f_u^T\lambda)_x \\ A_{12} &= -f_u[(f_u^T\lambda)_u]^{-1} f_u^T \\ A_{21} &= -(f_x^T\lambda)_x + (f_x^T\lambda)_u[(f_u^T\lambda)_u]^{-1}(f_u^T\lambda)_x \\ A_{22} &= -f_x^T + (f_x^T\lambda)_u[(f_u^T\lambda)_u]^{-1} f_u^T \\ C_1 &= f + (\hat{t} - t_0)\{f_{\hat{t}} - f_u[(f_u^T\lambda)_u]^{-1}(f_u^T\lambda)_{\hat{t}}\} \\ C_2 &= -f_x^T\lambda - (\hat{t} - t_0)\{(f_x^T\lambda)_{\hat{t}} - (f_x^T\lambda)_u[(f_u^T\lambda)_u]^{-1}(f_u^T\lambda)_{\hat{t}}\} \end{aligned} \tag{14}$$

and for $k = 1$

$$\begin{aligned} P_{11} &= g - f_u[(f_u^T\lambda)_u]^{-1} g_u^T\lambda \\ P_{21} &= -g_x^T\lambda - (f_x^T\lambda)_u[(f_u^T\lambda)_u]^{-1} g_u^T\lambda \end{aligned} \tag{15}$$

All the matrices in Eq. (14) are evaluated at the zero-order solution values. To complete the necessary conditions, it is also required to expand the boundary conditions and the transversality condition in Eq. (3). Note that Eq. (13) explicitly shows the effect of higher order corrections to the final time T. If the solution process is terminated at, say, $k = 1$, then a real-time sampled data implementation of the control solution would be constructed as follows. For the original system in Eq. (2), an expression for the optimal control is obtained as a function of x and λ from the optimality condition in Eq. (5). Then, treating the present state as the initial state, a first-order approximation is obtained by using $\lambda_0(t_0) + \epsilon\lambda_1(t_0)$ as an approximation for $\lambda(t_0)$ to compute the control, where $\lambda_0(t_0)$ and $\lambda_1(t_0)$ are obtained from the solutions of the zero- and first-order necessary conditions. This process is repeated at the next control update time by regarding the value of the state as the new initial state. Therefore, it is necessary to repeat the zero- and first-order solutions in updating the estimate of the costate variable.

The nonhomogeneous linear ordinary differential equations in Eqs. (13–15) may be expressed in terms of a convolution by first obtaining a state transition matrix. The state transition matrix $\Omega_A(\hat{t}, t_0)$ is merely the partial derivative of the zero-order solution at $\hat{t}$ with respect to the initial conditions $x_0(t_0)$ and $\lambda_0(t_0)$; hence it is easily computed given an analytic zero-order solution. In Appendix A it is shown that the result can be expressed in the form

$$\begin{aligned} \begin{bmatrix} x_k(\hat{t}) \\ \lambda_k(\hat{t}) \end{bmatrix} &= \Omega_A(\hat{t}, t_0)\begin{bmatrix} x_k(t_0) \\ \lambda_k(t_0) \end{bmatrix} + \int_{t_0}^{\hat{t}} \Omega_A(\hat{t}, \tau)\left\{\frac{T_k}{T_0}\begin{bmatrix} C_1(\tau) \\ C_2(\tau) \end{bmatrix}\begin{bmatrix} P_{1k}(\tau) \\ P_{2k}(\tau) \end{bmatrix}\right\} \mathrm{d}\tau \\ &= \Omega_A(\hat{t}, t_0)\begin{bmatrix} x_k(t_0) \\ \lambda_k(t_0) \end{bmatrix} + T_k\frac{\hat{t} - t_0}{T_0}\begin{bmatrix} \dot{x}_0(\hat{t}) \\ \dot{\lambda}_0(\hat{t}) \end{bmatrix} + \int_{t_0}^{\hat{t}} \Omega_A(\hat{t}, \tau)\left[\begin{bmatrix} P_{1k}(\tau) \\ P_{2k}(\tau) \end{bmatrix}\right] \mathrm{d}\tau \end{aligned} \tag{16}$$

Using the above expression at $\hat{t} = T_0$ along with the expansions of the boundary conditions, we can solve for $\lambda_k(t_0)$, v_k, and T_k from a set of linear algebraic equations.[5] Thus the major part of the computation for the first-order term lies in the quadrature that must be performed in Eq. (16). In a discrete time implementation, if the current state is regarded as the initial state, then $x_k(t_0) = 0$ in Eq. (16) since $x_0(t_0)$ satisfies the initial condition on the state variable. Since the zero-order solution changes at each update of the initial state, it is necessary to repeat the quadrature at each update for the higher order corrections. Alternatively, we can fix the zero-order solution and treat $x_1(t_0)$ as the deviation between the current state and the zero-order solution computed for the original epoch time but evaluated at the present time. In this form it would be possible to precompute the quadrature and store it as a function of a monotonic variable along the trajectory. Thus the real-time process of solving the zero-order problem and the quadrature can be avoided.

The case of discontinuous dynamics, such as might arise in a piecewise representation of the original problem, can be handled by a simple modification of Eq. (16). For example, in a two-stage representation we would have

$$\begin{bmatrix} x_k(\hat{t}) \\ \lambda_k(\hat{t}) \end{bmatrix} = \Omega_A^{(2)}(\hat{t}, t_s)\left\{\Omega_A^{(1)}(t_s, t_0)\begin{bmatrix} x_k(t_0) \\ \lambda_k(t_0) \end{bmatrix} + T_k\frac{t_s - t_0}{T_0}\begin{bmatrix} \dot{x}_0^{(1)}(t_s) \\ \dot{\lambda}_0^{(1)}(t_s) \end{bmatrix} + \int_{t_0}^{t_s} \Omega_A^{(1)}(\hat{t}, \tau)\begin{bmatrix} P_{1k}^{(1)}(\tau) \\ P_{1k}^{(2)}(\tau) \end{bmatrix} d\tau\right\} + T_k\frac{\hat{t} - t_s}{T_0}\begin{bmatrix} \dot{x}_0^{(2)}(\hat{t}) \\ \dot{\lambda}_0^{(2)}(\hat{t}) \end{bmatrix} + \int_{t_s}^{\hat{t}} \Omega_A^{(2)}(\hat{t}, \tau)\begin{bmatrix} P_{1k}^{(2)}(\tau) \\ P_{2k}^{(2)}(\tau) \end{bmatrix} d\tau, \qquad \hat{t} > t_s \tag{17}$$

The superscripts (1) and (2) denote the expressions for different sets of dynamics and t_s is the interior point where discontinuity occurs. This is needed in the next section when considering collocation methods that give rise to a discontinuity in first- or higher order derivatives at the nodes.

As stated earlier, the motivation behind a regular expansion of the solution lies in the assumption that the zero-order problem in Eq. (12) is considerably simpler to solve than the original problem in Eqs. (2–5). If numerical methods are still needed to solve the reduced problem, then little is gained by introducing regular expansions. Therefore, the reduced problem in a regular expansion should be analytically tractable to be of any value. On the other hand, one must retain enough of the original problem dynamics so that the zero-order solution is a reasonable approximation. These conflicting requirements may not be reconcilable in many applications of practical interest, unless a highly efficient numerical method is used to solve the zero-order problem. In the next section, we consider the use of a collocation method and artificially introduce a regular perturbation problem formulation whose solution serves to improve the collocation solution.

Method of Collocation

Collocation is a method for constructing an approximate solution to a set of differential equations by segmenting the time interval and representing the solution by piecewise polynomials. The unknown coefficients are determined by enforcing continuity at the nodes and that the derivatives of the interpolating functions satisfy the differential equations at some specified points within each segment. In this section we consider an optimization problem with unperturbed dynamics $dx/dt = f(x, u, t)$ and Hamiltonian $H = \lambda^T f$. For simplicity, assume a first-order polynomial approximation where the derivative constraints are enforced at the midpoint of each segment. These constraints can be expressed as

$$p_j = \frac{x_j - x_{j-1}}{\hat{t}_j - \hat{t}_{j-1}} = \left.\frac{\partial H}{\partial \lambda}\right|_{\hat{t}=(\hat{t}_j+\hat{t}_{j-1})/2;x=(x_j+x_{j-1})/2;\lambda=(\lambda_j+\lambda_{j-1})/2} \tag{18}$$

$$q_j = \frac{\lambda_j - \lambda_{j-1}}{\hat{t}_j - \hat{t}_{j-1}} = -\left.\frac{\partial H}{\partial x}\right|_{\hat{t}=(\hat{t}_j+\hat{t}_{j-1})/2;x=(x_j+x_{j-1})/2;\lambda=(\lambda_j+\lambda_{j-1})/2} \tag{19}$$

$$x(\hat{t}) = \begin{cases} x_{j-1} + p_j(\hat{t} - \hat{t}_{j-1}), & j = 1, \ldots, N \qquad (20) \\ \lambda_{j-1} + q_j(\hat{t} - \hat{t}_{j-1}), & \hat{t} \in [\hat{t}_{j-1}, \hat{t}_j], \\ & \hat{t}_0 = t_0, \hat{t}_N = t_0 + T_0 \qquad (21) \end{cases}$$

where N is the number of segments. The control is assumed to have been eliminated using the optimality condition. In practice, it is more convenient to directly evaluate the nodal values ($x_0, \lambda_0, \ldots, x_N, \lambda_N$) rather than finding the coefficients of the interpolating functions. Though higher order polynomials such as Hermite's cubic are generally preferred (because of their smoothness properties), we consider a first-order representation to simplify the presentation, although the approach applies equally well for higher order representations.

A regular perturbation formulation may be introduced by rewriting the actual dynamics in the following form:

$$\begin{aligned} \dot{x} &= p_j + \epsilon(H_\lambda - p_j) \\ \dot{\lambda} &= q_j + \epsilon(-H_x - q_j) \\ H_u &= 0, \qquad \hat{t} \in [\hat{t}_{j-1}, \hat{t}_j] \end{aligned} \tag{22}$$

Note that ϵ has been introduced as a bookkeeping parameter and has a nominal value of 1.0. The justification for this step is that if the collocation solution alone accurately approximates the true solution, then the second terms in Eq. (22) may be regarded as having a small perturbing effect on the state and costate derivatives, which actually is zero at the midpoints of the segments. If the control cannot be eliminated explicitly in the collocation formulation in Eqs. (18–21), then an analytic portion $\Pi(u, x, \lambda)$ of the optimality condition (for which it is possible to eliminate u) can be extracted such that

$$0 = \Pi + \epsilon(H_u - \Pi) \tag{23}$$

Note that in the above equations H is the Hamiltonian corresponding to the original system without a perturbation parameter. As presented above, a collocation solution may be viewed as the zero-order solution for the regular perturbation problem formulated in Eq. (22). Also, as will be shown by example in the next section, more intelligent choices of interpolation functions can be identified from the necessary conditions, by utilizing to the extent possible the analytically tractable portions of the solution. This results in a significant decrease in the computational requirements for a given level of accuracy.

Now we can apply the perturbation technique described in the previous section to improve the approximate zero-order solution from collocation. For the higher order problems defined in Eqs. (13–15) we have

$$A_{11_j} = \frac{\partial p_j}{\partial x} = \left.\frac{\partial^2 H}{\partial x \partial \lambda}\right|_{\hat{t}=(\hat{t}_j+\hat{t}_{j-1})/2;x=(x_j+x_{j-1})/2;\lambda=(\lambda_j+\lambda_{j-1})/2}$$

$$A_{12_j} = \frac{\partial p_j}{\partial \lambda} = \left.\frac{\partial^2 H}{\partial^2 \lambda}\right|_{\hat{t}=(\hat{t}_j+\hat{t}_{j-1})/2;x=(x_j+x_{j-1})/2;\lambda=(\lambda_j+\lambda_{j-1})/2}$$

$$A_{21_j} = \frac{\partial q_j}{\partial x} = -\left.\frac{\partial^2 H}{\partial^2 x}\right|_{\hat{t}=(\hat{t}_j+\hat{t}_{j-1})/2;x=(x_j+x_{j-1})/2;\lambda=(\lambda_j+\lambda_{j-1})/2}$$

$$A_{22_j} = \frac{\partial q_j}{\partial \lambda} = -\left.\frac{\partial^2 H}{\partial \lambda\, \partial x}\right|_{\hat{t}=(\hat{t}_j+\hat{t}_{j-1})/2;x=(x_j+x_{j-1})/2;\lambda=(\lambda_j+\lambda_{j-1})/2}$$

$$C_{1_j} = p_j + (\hat{t} - \hat{t}_{j-1})p_{t_j} = \left.\left(\frac{\partial H}{\partial \lambda} + (\hat{t} - \hat{t}_{j-1})\frac{\partial^2 H}{\partial t \partial \lambda}\right)\right|_{\hat{t} = (\hat{t}_j + \hat{t}_{j-1})/2;\, x = (x_j + x_{j-1})/2;\, \lambda = (\lambda_j + \lambda_{j-1})/2}$$

$$C_{2_j} = q_j + (\hat{t} - \hat{t}_{j-1})q_{t_j} = \left.\left(-\frac{\partial H}{\partial x} - (\hat{t} - \hat{t}_{j-1})\frac{\partial^2 H}{\partial t\, \partial x}\right)\right|_{\hat{t} = (\hat{t}_j + \hat{t}_{j-1})/2;\, x = (x_j + x_{j-1})/2;\, \lambda = (\lambda_j + \lambda_{j-1})/2} \tag{24}$$

and for $k = 1$,

$$\begin{aligned} P_{11_j} &= \left.\frac{\partial H}{\partial \lambda}\right|_{\hat{t};x=x_{j-1}+p_j(\hat{t}-\hat{t}_{j-1});\lambda=\lambda_{j-1}+q_j(\hat{t}-\hat{t}_{j-1})} - p_j \\ P_{21_j} &= -\left.\frac{\partial H}{\partial x}\right|_{\hat{t};x=x_{j-1}+p_j(\hat{t}-\hat{t}_{j-1});\lambda=\lambda_{j-1}+q_j(\hat{t}-\hat{t}_{j-1})} - q_j \end{aligned} \tag{25}$$

where all the terms in Eq. (24) are constant within a segment and are evaluated using the collocation solution. The matrix A_j is simply the perturbation of the original state and costate dynamics evaluated at the constraint point of each segment. The expression in Eq. (16), which now corresponds to a piecewise constant system matrix A_j, can be written as

$$\begin{bmatrix} x_k(\hat{t}) \\ \lambda_k(\hat{t}) \end{bmatrix} = \Omega_{A_j}(\hat{t}, \hat{t}_{j-1}) \begin{bmatrix} x_k(\hat{t}_{j-1}) \\ \lambda_k(\hat{t}_{j-1}) \end{bmatrix} + \frac{T_k}{T_0} \Omega_{A_j}(\hat{t}, \hat{t}_{j-1}) A_j^{-1}$$
$$\times \left(\begin{bmatrix} p_j \\ q_j \end{bmatrix} + (\hat{t} - \hat{t}_{j-1}) \begin{bmatrix} p_{t_j} \\ q_{t_j} \end{bmatrix} + A_j^{-1} \begin{bmatrix} p_{t_j} \\ q_{t_j} \end{bmatrix} \right)$$
$$+ \int_{t_{j-1}}^{\hat{t}} \Omega_{A_j}(\hat{t}, \tau) \begin{bmatrix} P_{1k}(\tau) \\ P_{2k}(\tau) \end{bmatrix} \mathrm{d}\tau, \quad \hat{t} \in [\hat{t}_j, \hat{t}_{j-1}] \tag{26}$$

where Ω_{A_j} is the state transition matrix and p_{t_j}, q_{t_j} are defined in Eq. (24). Note that Ω_{A_j} is not the same as in Eqs. (16) and (17) because A is defined differently. The state transition matrix here may not have an analytic expression because the zero-order solution is not necessarily analytic. If this is true, we can solve Eqs. (24) and (25) using the sensitivity functions and superposition property of linear systems. This is done by assigning a unit vector for the initial conditions and numerically integrating the system from t_0 to $t_0 + T_0$. Thus by changing the position of the nonzero element in the unit vectors, the sensitivity functions are obtained. This process can be done in parallel for different unit vectors.

In the zero-order solution, ϵ in Eq. (22) is set to zero, which means that the standard collocation constraints in Eqs. (18–21) are employed and an approximate solution is obtained by solving the algebraic equations. Then first- and higher order corrections may be computed by quadrature, as explained in the preceding section on regular perturbation.

Duffing Equation Example

This investigation was carried out to demonstrate the hybrid approach outlined in the preceding section. The example is based on Duffing's equation presented in its first-order form:

$$\begin{aligned} \dot{x} &= v \qquad & x(0) &= x_0 \\ \dot{v} &= -x - ax^3 + u, \qquad & v(0) &= v_0 \end{aligned} \tag{27}$$

and the objective is

$$\min_u \left[S_x x^2(t_f) + S_v v^2(t_f) + \int_0^{t_f} \left(1 + \frac{u^2}{2}\right) \mathrm{d}t \right] \tag{28}$$

with S_x, S_v being the weights on the terminal values and t_f is free. The problem can be converted to the Mayer form in Eq. (1) (if desired) through the usual method of introducing an additional state equation whose right-hand side is the integrand of Eq. (28). We investigate the problem in different levels of complexity according to how the dynamics of the full system are treated.

Level 0 Formulation

This is the degenerate case in which there is an analytic zero-order solution, and therefore collocation is not required. Let $\epsilon = a$, thus neglecting the hardening effect ax^3 in the original problem. The necessary conditions are

$$\begin{aligned} \dot{x} &= v, \qquad & x(0) &= x_0 \\ \dot{v} &= -x + u - \epsilon x^3, \qquad & v(0) &= v_0 \\ \dot{\lambda}_x &= \lambda_v + \epsilon 3\lambda_v x^2, \qquad & \lambda_x(t_f) &= 2S_x x(t_f) \\ \dot{\lambda}_v &= -\lambda_x, \qquad & \lambda_v(t_f) &= 2S_v v(t_f) \end{aligned} \tag{29}$$
$$H_u = u + \lambda_v = 0$$
$$\{H = \lambda_x v + \lambda_v(-x + u - \epsilon x^3) + 1 + \tfrac{1}{2}u^2\}\big|_{t_f} = 0$$

The zero-order problem ($\epsilon = 0$) is linear and time-invariant and can easily be solved as

$$\begin{bmatrix} x_0(\hat{t}) \\ v_0(\hat{t}) \\ \lambda_{x0}(\hat{t}) \\ \lambda_{v0}(\hat{t}) \end{bmatrix} = \begin{bmatrix} \cos\bar{t} & \sin\bar{t} & \frac{1}{2}(\sin\bar{t} - \bar{t}\cos\bar{t}) & -\frac{1}{2}\bar{t}\sin\bar{t} \\ -\sin\bar{t} & \cos\bar{t} & \frac{1}{2}\bar{t}\sin\bar{t} & -\frac{1}{2}(\sin\bar{t} + \bar{t}\cos\bar{t}) \\ 0 & 0 & \cos\bar{t} & \sin\bar{t} \\ 0 & 0 & -\sin\bar{t} & \cos\bar{t} \end{bmatrix} \times \begin{bmatrix} x_0(\hat{t}_0) \\ v_0(\hat{t}_0) \\ \lambda_{x0}(\hat{t}_0) \\ \lambda_{v0}(\hat{t}_0) \end{bmatrix} \tag{30}$$

where

$$\bar{t} = \hat{t} - \hat{t}_0, \qquad \hat{t}_0, \hat{t} \in [0, T_0] \tag{31}$$

The above state transition matrix is also the state transition matrix $\Omega_A(\hat{t}, \hat{t}_0)$ for the higher order correction. Given the boundary conditions $x_0(0) = x_0$, $v_0(0) = v_0$, $\lambda_{x0}(T_0) = 2S_x x_0(T_0)$, $\lambda_{v0}(T_0) = 2S_v v_0(T_0)$, the remaining unknowns $\lambda_{x0}(0)$, $\lambda_{v0}(0)$, $\lambda_{x0}(T_0)$, $\lambda_{v0}(T_0)$, T_0 can be solved with Newton's method using Eq. (30) and the corresponding transversality condition

$$\{H_0 = \lambda_{x0} v_0 - \lambda_{v0} x_0 - \tfrac{1}{2}\lambda_{v0}^2 + 1\}\big|_{T_0} = 0 \tag{32}$$

From Eq. (13), the differential equations for the higher order problems are as follows:

$$\frac{\mathrm{d}}{\mathrm{d}\hat{t}} \begin{bmatrix} x_k \\ v_k \\ \lambda_{xk} \\ \lambda_{vk} \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & -1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{bmatrix} \begin{bmatrix} x_k \\ v_k \\ \lambda_{xk} \\ \lambda_{vk} \end{bmatrix} + \frac{T_k}{T_0} \begin{bmatrix} v_0(\hat{t}) \\ -x_0(\hat{t}) - \lambda_{v0}(\hat{t}) \\ \lambda_{v0}(\hat{t}) \\ -\lambda_{x0}(\hat{t}) \end{bmatrix} + \begin{bmatrix} P_{1k}(\hat{t}) \\ P_{2k}(\hat{t}) \\ P_{3k}(\hat{t}) \\ P_{4k}(\hat{t}) \end{bmatrix} \tag{33}$$

with the boundary conditions

$$x_k(0) = v_k(0) = 0, \qquad \lambda_{xk}(T_0) = 2S_x x_k(T_0)$$
$$\lambda_{vk}(T_0) = 2S_v v_k(T_0) \tag{34}$$

In this case, we have, for $k = 1, 2$,

$$P_{11} = 0, \qquad P_{21} = -x_0^3, \qquad P_{31} = 3\lambda_{v0}x_0^2, \qquad P_{41} = 0$$
$$P_{12} = v_1 T_1 / T_0,$$
$$P_{22} = -(x_1 + \lambda_{v1} + x_0^3) T_1 / T_0 - 3x_0^2 x_1 \tag{35}$$
$$P_{32} = (\lambda_{v1} + 3\lambda_{v0}x_0^2) T_1 / T_0 + 3\lambda_{v1}x_0^2 + 3\lambda_{v0}x_0 x_1$$
$$P_{42} = -\lambda_{x1} T_1 / T_0$$

and the transversality conditions

$$\{H_1 = \lambda_{x1} v_0 - \lambda_{v1}(x_0 + \lambda_{v0}) + \lambda_{x0} v_1 - \lambda_{v0}(x_1 + x_0^3)\}|_{T_0} = 0 \tag{36}$$

$$\{H_2 = \lambda_{x2} v_0 - \lambda_{v2}(x_0 + \lambda_{v0}) + \lambda_{x0} v_2 + \lambda_{v1} x_1 - 3\lambda_{v0} x_0^2 x_1 - \lambda_{v1}(x_1 + \lambda_{v1} + x_0^3) + \tfrac{1}{2}\lambda_{v1}^2\}|_{T_0} = 0 \tag{37}$$

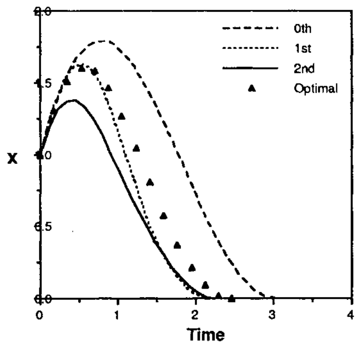

Fig. 1 Level 0 result for $x(t)$.

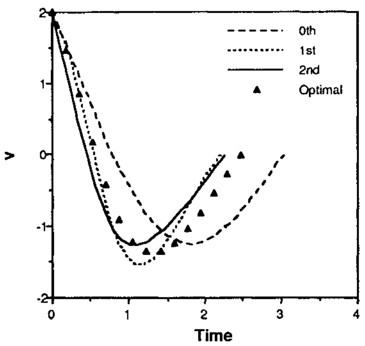

Fig. 2 Level 0 result for $v(t)$.

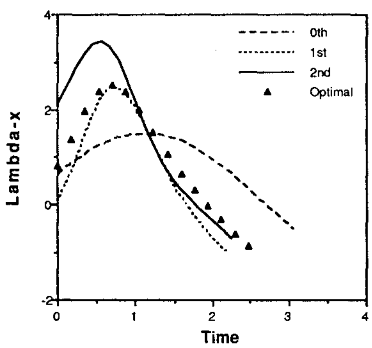

Fig. 3 Level 0 result for $\lambda_x(t)$.

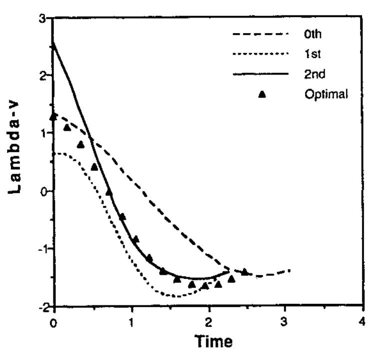

Fig. 4 Level 0 result for $\lambda_v(t)$.

which are needed to compute the first and second-order corrections by quadrature. The results are shown in Figs. 1–4 for $x_0 = 1$, $v_0 = 2$, $S_x = S_v = 100$, and $a = 0.4$. The first-order state and costate histories are stored and later retrieved by linear interpolation to construct the second-order solution. The optimal solution generated using a multiple shooting technique[10] is also included for comparison. These results clearly show that the series is not convergent and that the most accurate approximation is obtained using a first-order solution. If we regard this level of accuracy as insufficient, then the conclusion must be that the nonlinear term ax^3 is too large to be neglected in the zero-order solution.

Level 1 Formulation

This case illustrates the hybrid approach as outlined in the section on collocation using a piecewise linear representation to approximate the states and the costates for the zero-order solution. The dynamics in Eq. (29) are first written in the form of Eq. (22). This entails substituting $\epsilon = a$ (its actual value) and introducing (artificially) a newly defined ϵ as the bookkeeping parameter for the expansion. The interpolatory constraints for an N equally spaced segmentation are

$$\frac{x_{0j} - x_{0j-1}}{T_0/N} = \frac{v_{0j} + v_{0j-1}}{2} = p_{xj} \tag{38}$$

$$\frac{v_{0j} - v_{0j-1}}{T_0/N} = -\frac{x_{0j} + x_{0j-1}}{2} - \frac{\lambda_{v0j} + \lambda_{v0j-1}}{2} - a\left(\frac{x_{0j} + x_{0j-1}}{2}\right)^3 = p_{vj} \tag{39}$$

$$\frac{\lambda_{x0j} - \lambda_{x0j-1}}{T_0/N} = \frac{\lambda_{v0j} + \lambda_{v0j-1}}{2} \times \left[1 + 3a\left(\frac{x_{0j} + x_{0j-1}}{2}\right)^2\right] = q_{xj} \tag{40}$$

$$\frac{\lambda_{v0j} - \lambda_{v0j-1}}{T_0/N} = -\frac{\lambda_{x0j} - \lambda_{x0j-1}}{2} = q_{vj} \tag{41}$$

with the boundary conditions and transversality condition given by

$$x_{00} = x_0, \qquad v_{00} = v_0, \qquad \lambda_{x0N} = 2S_x x_{0N}$$
$$\lambda_{v0N} = 2S_v v_{0N} \tag{42}$$
$$\lambda_{x0N} v_{0N} + \lambda_{v0N}(-x_{0N} - \lambda_{v0N} - a x_{0N}^3) + \tfrac{1}{2}\lambda_{v0N}^2 + 1 = 0$$

There are $4N + 5$ equations to solve for the $4N + 5$ unknowns of $x_{00}, v_{00}, \lambda_{x00}, \lambda_{x00}, \ldots, x_{0N}, v_{0N}, \lambda_{x0N}, \lambda_{v0N}, T_0$. Solutions for several values of N are presented in Figs. 5–8. Note that accuracy improves with increasing N, but at the expense of having to solve a large nonlinear system of equations.

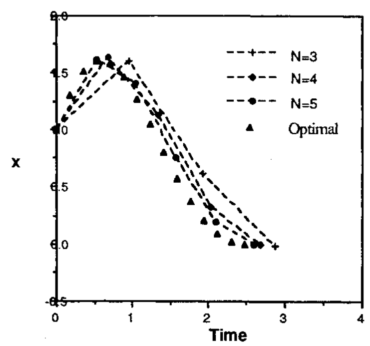

Fig. 5 Level 1 zero-order results for $x(t)$ for different values of N.

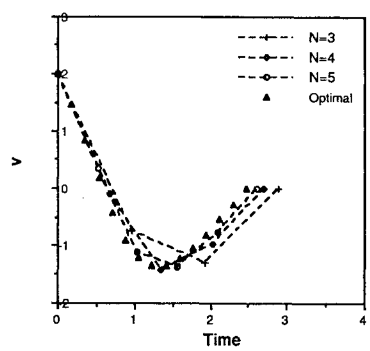

Fig. 6 Level 1 zero-order results for $v(t)$ for different values of N.

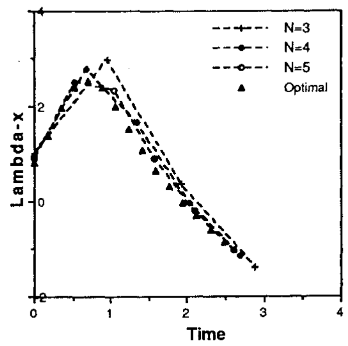

Fig. 7 Level 1 zero-order results for $\lambda_x(t)$ for different values of N.

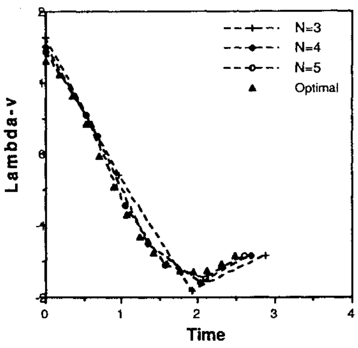

Fig. 8 Level 1 zero-order results for $\lambda_v(t)$ for different values of N.

The higher order dynamics in this case are

$$\frac{\mathrm{d}}{\mathrm{d}\hat{t}}\begin{bmatrix} x_k \\ v_k \\ \lambda_{xk} \\ \lambda_{vk} \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ -c & 0 & 0 & -1 \\ b & 0 & 0 & c \\ 0 & 0 & -1 & 0 \end{bmatrix}\begin{bmatrix} x_k \\ v_k \\ \lambda_{xk} \\ \lambda_{vk} \end{bmatrix} + \frac{T_k}{T_0}\begin{bmatrix} p_{xj} \\ p_{vj} \\ q_{xj} \\ q_{vj} \end{bmatrix} + \begin{bmatrix} P_{1k}(\hat{t}) \\ P_{2k}(\hat{t}) \\ P_{3k}(\hat{t}) \\ P_{4k}(\hat{t}) \end{bmatrix}, \qquad \hat{t} \in [\hat{t}_j, \hat{t}_{j-1}] \tag{43}$$

where

$$c = 1 + 3ax_0^2|_{\bar{t}}, \qquad b = 6a(\lambda_{v0}x_0)|_{\bar{t}}, \qquad \bar{t} = \tfrac{1}{2}(\hat{t}_j + \hat{t}_{j-1}) \tag{44}$$

The state transition matrix expression for this case is given in Appendix B. For $k = 1, 2$, the forcing function terms in Eq. (43) are

$$\begin{aligned} P_{11} &= v_0 - p_{xj}, \qquad P_{21} = -x_0 - \lambda_{v0} - ax_0^3 - p_{vj} \\ P_{31} &= \lambda_{v0}(1 + 3ax_0^2) - q_{xj}, \qquad P_{41} = -\lambda_{x0} - q_{vj} \\ P_{12} &= (v_1 + v_0 - p_{xj})T_1/T_0 \\ P_{22} &= (-cx_1 - \lambda_{v1} - x_0 - \lambda_{v0} - ax_0^3 - p_{vj}) \\ &\quad \times T_1/T_0 - 3ax_0|_{\bar{t}}x_1^2 - (1 + 3ax_0^2 - c)x_1 \\ P_{32} &= [c\lambda_{v1} + bx_1 + \lambda_{v0}(1 + 3ax_0^2 - q_{xj})] \\ &\quad \times T_1/T_0 + 6ax_0\big|_{\bar{t}}\lambda_{v1}x_1 + 3a\lambda_{v0}|_{\bar{t}}x_1^2 \\ &\quad + (1 + 3ax_0^2 - c)\lambda_{v1} + (6a\lambda_{v0}x_0 - b)x_1 \\ P_{42} &= (-\lambda_{x1} - \lambda_{x0} - q_{vj})T_1/T_0 \end{aligned} \tag{45}$$

where

$$\begin{aligned} x_0(\hat{t}) &= x_{0j-1} + p_{xj}(\hat{t} - \hat{t}_{j-1}) \\ v_0(\hat{t}) &= v_{0j-1} + p_{vj}(\hat{t} - \hat{t}_{j-1}) \\ \lambda_{x0}(\hat{t}) &= \lambda_{x0j-1} + q_{xj}(\hat{t} - \hat{t}_{j-1}) \\ \lambda_{v0}(\hat{t}) &= \lambda_{v0j-1} + q_{vj}(\hat{t} - \hat{t}_{j-1}) \end{aligned} \tag{46}$$

The boundary conditions in Eq. (34) are applied by replacing $x_k(0)$, $v_k(0)$, $x_k(T_0)$, $v_k(T_0)$, $\lambda_{xk}(T_0)$, $\lambda_{vk}(T_0)$ with x_{k0}, v_{k0}, x_{kN}, v_{kN}, λ_{xkN}, λ_{vkN}. The corresponding expansions of the transversality conditions in this case are

$$\begin{aligned} 0 &= \lambda_{x1N}v_{0N} + \lambda_{x0N}v_{1N} - \lambda_{v1N}(x_{0N} + ax_{0N}^3) \\ &\quad + \lambda_{v0N}(-x_{1N} - \lambda_{v0N} - ax_{0N}^3) \\ 0 &= \lambda_{x2N}v_{0N} + \lambda_{x0N}v_{2N} + \lambda_{x1N}v_{1N} - \lambda_{v2N}(x_{0N} + ax_{0N}^3) \\ &\quad + \lambda_{v0N}(-x_{2N} - \lambda_{v2N} - 3ax_{0N}^2x_{2N}) \\ &\quad + \lambda_{v1N}(-x_{1N} - 3ax_{0N}^2x_1) - \tfrac{1}{2}\lambda_{v1N}^2 \end{aligned} \tag{47}$$

First- and second-order corrections were computed for the case where $N = 3$ was used in the zero-order collocation solution. The results are shown in Figs. 9–12. Comparison with the $N = 3$ results in Figs. 5–8 shows that a significant improvement in accuracy is achievable without requiring a large number of segments. In Figs. 9–12 the second-order solution is indistinguishable from the optimal solution. The discontinuity in slope (which is a consequence of using first-order interpolation functions for the collocation solution) is also smoothed as the order of the correction increases. Contrary to the level 0 formulation results, the second-order corrections do not diverge due to the fact that the nonlinear term has been accounted for in the zero-order collocation solution.

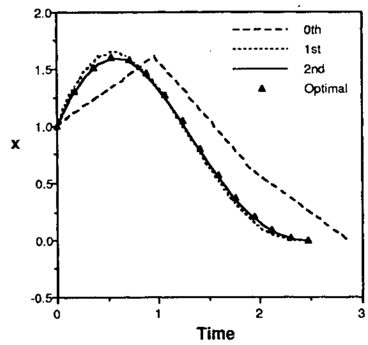

Fig. 9 Level 1 higher order results for $x(t)$ for $N = 3$.

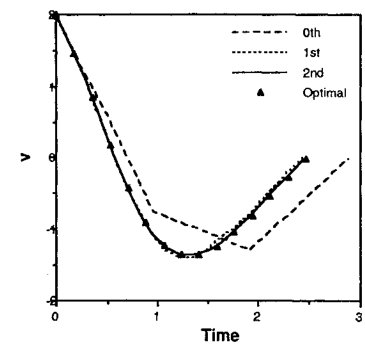

Fig. 10 Level 1 higher order results for $v(t)$ for $N = 3$.

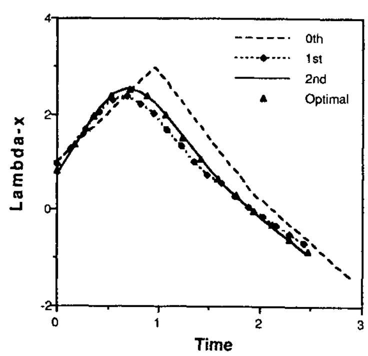

Fig. 11 Level 1 higher order results for $\lambda_x(t)$ for $N = 3$.

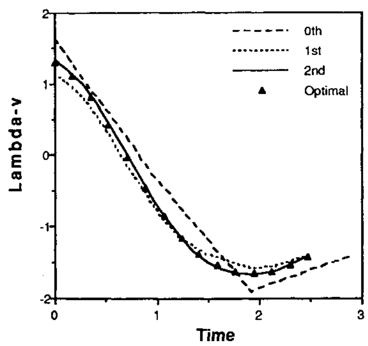

Fig. 12 Level 1 higher order results for $\lambda_v(t)$ for $N = 3$.

Level 2 Formulation

As a second illustration of a hybrid solution approach we retain a portion of the dynamics from the necessary conditions to identify a more intelligent interpolating function for the hybrid level 1 formulation. Consider the following simple modification of Eq. (22) for this example:

$$\begin{aligned}\dot{x} &= v\\ \dot{v} &= p_{vj} + \epsilon(-x - \lambda_v - ax^3 - p_{vj})\\ \dot{\lambda}_x &= q_{xj} + \epsilon[\lambda_v(1 + 3ax^2) - q_{xj}]\\ \dot{\lambda}_v &= -\lambda_x\end{aligned} \tag{48}$$

Note that now we interpolate only the variables that have nonlinear coupling (v and λ_x) and that the interpolations for x and λ_x retain the dynamics in the original problem. This portion is analytically tractible and therefore should provide a more suitable interpolation than the simple first-order forms assumed in the level 0 formulation. The interpolating functions in this case are

$$x_0(\hat{t}) = x_0(\hat{t}_{j-1}) + [v_{0j-1} + \tfrac{1}{2}p_{vj}(\hat{t} - \hat{t}_{j-1})](\hat{t} - \hat{t}_{j-1}) \tag{49}$$

$$v_0(\hat{t}) = v_0(\hat{t}_{j-1}) + p_{vj}(\hat{t} - \hat{t}_{j-1}) \tag{50}$$

$$\lambda_{x0}(\hat{t}) = \lambda_{x0j-1} + q_{xj}(\hat{t} - \hat{t}_{j-1}) \tag{51}$$

$$\lambda_{v0}(\hat{t}) = \lambda_{v0}(\hat{t}_{j-1}) - [\lambda_{x0j-1} + \tfrac{1}{2}q_{xj}(\hat{t} - \hat{t}_{j-1})](\hat{t} - \hat{t}_{j-1}) \tag{52}$$

Consequently, there are fewer unknowns ($2N + 5$) to be solved and the dynamics retained in the formulation should improve the zero-order collocation approximation. This allows even fewer segments to be used. To evaluate the zero-order solution, conditions in Eq. (42) are enforced by replacing x_{0N}, λ_{v0N} with $x_0(\hat{t}_N)$, $\lambda_{v0}(\hat{t}_N)$ from Eqs. (49) and (52), and similarly for the first-order expressions. The forcing terms for this case are

$$\begin{aligned}P_{11} &= 0, & P_{21} &= -x_0 - \lambda_{v0} - ax_0^3 - p_{vj}\\ P_{31} &= \lambda_{v0}(1 + 3ax_0^2) - q_{xj}, & P_{41} &= 0\end{aligned} \tag{53}$$

and the state transition matrix is the same as that in level 1.

Figures 13 and 14 show the zero- and first-order state solutions for the case $N = 2$. The results show that the zero-order solution is dramatically improved in comparison to the zero-order solution for $N = 3$ of the level 1 formulation in Figs. 9 and 10. The accuracy of the first-order solutions in Figs. 13 and 14 is very good and is almost riding on the exact solutions, even though a cruder segmentation has been used. A similar trend was observed for the costate histories.

Level 3 Formulation

In this last demonstration, the level 2 formulation is further enhanced. All the linear terms are retained in the zero-order solution, and the nonlinear terms in the v and λ_x dynamics are approximated by piecewise constants. The resultant expression is as follows:

$$\begin{aligned}\dot{x} &= v\\ \dot{v} &= -x - \lambda_v + p_{vj} + \epsilon(-ax^3 - p_{vj})\\ \dot{\lambda}_x &= \lambda_v + q_{xj} + \epsilon(\lambda_v 3ax^2 - q_{xj})\\ \dot{\lambda}_v &= -\lambda_x\end{aligned} \tag{54}$$

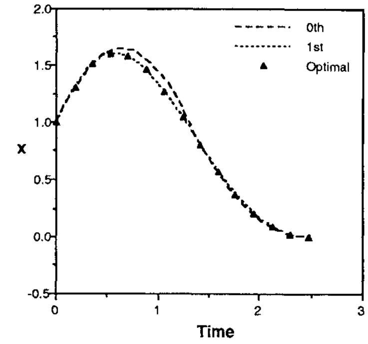

Fig. 13 Level 2 higher order results for $x(t)$ for $N = 2$.

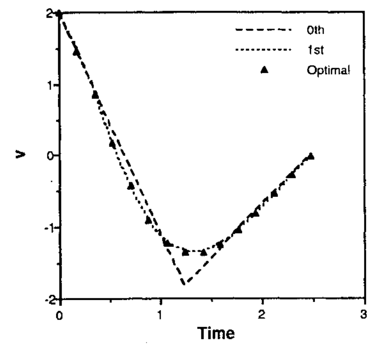

Fig. 14 Level 2 higher order results for $v(t)$ for $N = 2$.

This is equivalent to the level 0 formulation except for the interpolation constants p_{vj} and q_{xj}. This formulation represents an attempt to make maximum utilization of the analytically tractable portion of the solution in selecting the interpolating function for the collocation solution in the zero-order problem and therefore represents an improvement over the level 1 and 2 formulations. In comparison to the zero-order level 0 formulation, the constants p_{vj} and q_{xj} provide the additional freedom needed to satisfy the exact equations at the midpoints through the constraints imposed by Eqs. (18) and (19). Recall that in these equations H is the exact Hamiltonian for the original problem, whereas in the level 0 formulation the hardening effect is ignored in the zero-order solution. Hence, the effect of the nonlinear terms in Eq. (29) is partially accounted for in the zero-order solution. This also occurs in the level 1 and level 2 formulations.

The zero-order solutions in this case are similar to that for the level 0 case:

$$\begin{aligned}x_0(\hat{t}) &= [x_0(\hat{t}_{j-1}) - p_{vj} - q_{xj}]\cos\bar{t} + v_0(\hat{t}_{j-1})\sin t\\ &\quad + \lambda_{x0}(\hat{t}_{j-1})[\sin\bar{t} - \bar{t}\cos t]/2\\ &\quad - [\lambda_{v0}(\hat{t}_{j-1}) + q_{xj}]\bar{t}\sin t/2 + p_{vj} + q_{xj}\\ v_0(\hat{t}) &= -[x_0(\hat{t}_{j-1} - p_{vj} - q_{xj}]\sin\bar{t} + v_0(\hat{t}_{j-1})\cos\bar{t}\\ &\quad + \lambda_{x0}(\hat{t}_{j-1})\bar{t}\sin\bar{t}/2\\ &\quad - [\lambda_{v0}(\hat{t}_{j-1}) + q_{xj}](\sin\bar{t} + \bar{t}\cos\bar{t})/2\\ \lambda_{x0}(\hat{t}) &= \lambda_{x0}(\hat{t}_{j-1})\cos\bar{t} + [\lambda_{v0}(\hat{t}_{j-1}) + q_{xj}]\sin\bar{t}\\ \lambda_{v0}(\hat{t}) &= -\lambda_{x0}(\hat{t}_{j-1})\sin\bar{t} + [\lambda_{v0}(\hat{t}_{j-1}) + q_{xj}]\cos\bar{t} - q_{xj}\end{aligned} \tag{55}$$

where

$$\begin{aligned}p_{vj} &= -ax^3|_{x[(\hat{t}_j + \hat{t}_{j-1})/2]}\\ q_{xj} &= 3a(\lambda_v x^2)|_{\lambda_v[(\hat{t}_j + \hat{t}_{j-1})/2];x[(\hat{t}_j + \hat{t}_{j-1})/2]}\\ \bar{t} &= \hat{t} - \hat{t}_{j-1} \qquad \hat{t} \in [t_{j-1}, t_j]\end{aligned} \tag{56}$$

In this formulation, an efficient way to find the collocation solution is to solve for the $2N + 5$ unknowns of

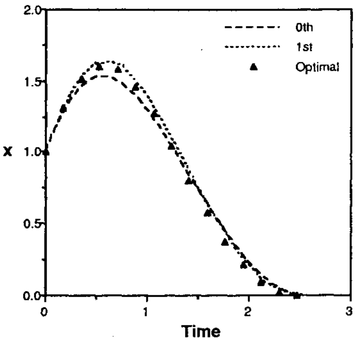

Fig. 15 Level 3 higher order results for $x(t)$ for $N = 1$.

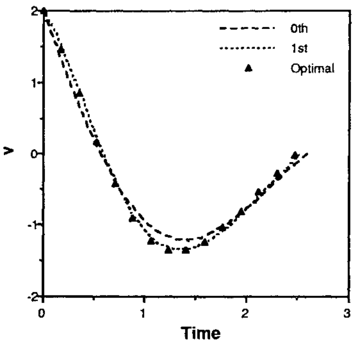

Fig. 16 Level 3 higher order results for $v(t)$ for $N = 1$.

$x_0(0), v_0(0), \lambda_{x0}(0), \lambda_{v0}(0), p_{v1}, q_{x1}, \ldots, p_{vN}, q_{xN}, t_{f0}$ using Eq. (55) in Eqs. (56) and (42). The high-order formulations are obtained in the same manner as the previous levels and are not repeated here. The zero- and first-order results using only one segment are shown in Figs. 15 and 16. Though the first-order results are not as accurate as those in level 2 (because only one segment is used), both zero- and first-order solutions are far superior than the level 0 results (Figs. 1 and 2), which correspond to the degenerate case of only one segment.

Conclusions

A hybrid analytic/numerical approach for solving optimization problems using regular perturbation and collocation methods has been presented. The hybrid approach shows that it is possible to significantly improve a collocation solution without increasing the number segments. The loss in accuracy that results from using a smaller number of segments is compensated by the addition of higher order corrections to the solution based on regular perturbation theory. Viewed a second way, using collocation to solve the zero-order problem in a regular perturbation expansion allows more of the dynamics to be retained in the zero-order solution. It has also been shown that further dramatic improvements can be achieved by selecting more intelligent interpolating functions that are derived from the analytically tractable portions of the necessary conditions. The approach has important implications in real-time guidance applications.

Appendix A

We want to show that

$$\frac{T_k}{T_0}\int_{t_0}^{\hat{t}} \Omega_A(\hat{t},\tau)\begin{bmatrix} C_1(\tau) \\ C_2(\tau)\end{bmatrix} \mathrm{d}\tau = T_k\frac{\hat{t}-t_0}{T_0}\begin{bmatrix}\dot{x}_0(\hat{t}) \\ \dot{\lambda}_0(\hat{t})\end{bmatrix} \tag{A1}$$

in Eq. 16. Let $f_1 = \dot{x}_0(x_0, \lambda_0, \tau)$, $f_2 = \dot{\lambda}_0(x_0, \lambda_0, \tau)$, assuming u being eliminated, and recall that

$$C_1 = f_1 + (\tau - t_0)\frac{\partial f_1}{\partial \tau}, \qquad C_2 = f_2 + (\tau - t_0)\frac{\partial f_2}{\partial \tau} \tag{A2}$$

The left-hand side of Eq. (A1) becomes

$$\frac{T_k}{T_0}\int_{t_0}^{\hat{t}} \Omega_A(\hat{t},\tau)\left\{\frac{\mathrm{d}}{\mathrm{d}\tau}\left((\tau-t_0)\begin{bmatrix} f_1 \\ f_2\end{bmatrix}\right) - (\tau - t_0)\right.$$
$$\left.\times\begin{bmatrix}\dfrac{\partial f_1}{\partial x_0} & \dfrac{\partial f_1}{\partial \lambda_0} \\ \dfrac{\partial f_2}{\partial x_0} & \dfrac{\partial f_2}{\partial \lambda_0}\end{bmatrix}\begin{bmatrix} f_1 \\ f_2\end{bmatrix}\right\}\mathrm{d}t \tag{A3}$$

Using integration by parts on the first term in Eq. (A3), we have

$$\frac{T_k}{T_0}\left((\tau-t_0)\Omega_A(\hat{t},\tau)\begin{bmatrix} f_1 \\ f_2\end{bmatrix}\right)\Bigg|_{\tau=t_0}^{\tau=\hat{t}}$$
$$-\frac{T_k}{T_0}\int_{t_0}^{\hat{t}}(\tau-t_0)\left(\frac{\mathrm{d}}{\mathrm{d}\tau}\Omega_A(\hat{t},\tau)\right)\begin{bmatrix} f_1 \\ f_2\end{bmatrix}\mathrm{d}\tau$$
$$-\frac{T_k}{T_0}\int_{t_0}^{\hat{t}}(\tau-t_0)\Omega_A(\hat{t},\tau)\begin{bmatrix}\dfrac{\partial f_1}{\partial x_0} & \dfrac{\partial f_1}{\partial \lambda_0} \\ \dfrac{\partial f_2}{\partial x_0} & \dfrac{\partial f_2}{\partial \lambda_0}\end{bmatrix}\begin{bmatrix} f_1 \\ f_2\end{bmatrix}\mathrm{d}\tau \tag{A4}$$

Substituting the state transition matrix property

$$\frac{\mathrm{d}}{\mathrm{d}\tau}\Omega_A(\hat{t},\tau) = -\Omega_A(\hat{t},\tau)\begin{bmatrix}\dfrac{\partial f_1}{\partial x_0} & \dfrac{\partial f_1}{\partial \lambda_0} \\ \dfrac{\partial f_2}{\partial x_0} & \dfrac{\partial f_2}{\partial \lambda_0}\end{bmatrix} \tag{A5}$$

into Eq. (A4), the last two terms cancel and the result is demonstrated. This result was also used in Eq. (26).

Appendix B

The state transition matrix for the level 1 case of the Duffing example is

$$\Omega_A(\hat{t}, t_0) = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{11} & -a_{14} & a_{24} \\ -ba_{24} & ba_{23} & a_{11} & -a_{21} \\ -ba_{23} & -ba_{13} & -a_{12} & a_{11}\end{bmatrix} \tag{B1}$$

For $b > 0$

$$\alpha = c - \sqrt{b}, \qquad \beta = c + \sqrt{b}, \qquad \bar{t} = \hat{t} - t_0$$

$$a_{11} = \frac{\alpha - c}{\alpha - \beta}\cos(\bar{t}\sqrt{\alpha}) + \frac{c-\beta}{\alpha-\beta}\cos(\bar{t}\sqrt{\beta})$$

$$a_{12} = \frac{\alpha - c}{(\alpha - \beta)\sqrt{\alpha}}\sin(\bar{t}\sqrt{\alpha}) + \frac{c-\beta}{(\alpha-\beta)\sqrt{\beta}}\sin(\bar{t}\sqrt{\beta})$$

$$a_{13} = \frac{-1}{(\alpha - \beta)\sqrt{\alpha}}\sin(\bar{t}\sqrt{\alpha}) + \frac{1}{(\alpha-\beta)\sqrt{\beta}}\sin(\bar{t}\sqrt{\beta}) \tag{B2}$$

$$a_{14} = \frac{1}{\alpha - \beta}\cos(\bar{t}\sqrt{\alpha}) - \frac{1}{\alpha-\beta}\cos(\bar{t}\sqrt{\beta})$$

$$a_{21} = \frac{c^2 - \alpha c - b}{(\alpha - \beta)\sqrt{\alpha}}\sin(\bar{t}\sqrt{\alpha}) + \frac{b + c\beta - c^2}{(\alpha-\beta)\sqrt{\beta}}\sin(\bar{t}\sqrt{\beta})$$

$$a_{24} = \frac{-\sqrt{\alpha}}{\alpha - \beta}\sin(\bar{t}\sqrt{\alpha}) + \frac{\sqrt{\beta}}{\alpha-\beta}\sin(\bar{t}\sqrt{\beta})$$

For $b < 0$

$$\theta = \frac{1}{2}\left(\sqrt{2\sqrt{c^2-b} - 2c}\right), \qquad \phi = \sqrt{\tfrac{1}{2}(c + \sqrt{c^2 - b})}$$

$$a_{12} = \frac{\theta^2 + \phi^2 - c}{2\theta(\theta^2 + \phi^2)} \sinh(\theta\bar{t}) \cos(\phi\bar{t}) + \frac{\theta^2 + \phi^2 + c}{2\phi(\theta^2 + \phi^2)} \cosh(\theta\bar{t}) \sin(\phi\bar{t})$$

$$a_{13} = \frac{-1}{2\theta(\theta^2 + \phi^2)} \sinh(\theta\bar{t}) \cos(\phi\bar{t}) + \frac{1}{2\phi(\theta^2 + \phi^2)} \cosh(\theta\bar{t}) \sin(\phi\bar{t}) \tag{B3}$$

$$a_{21} = \frac{c^2 - b - c(\theta^2 + \phi^2)}{2\theta(\theta^2 + \phi^2)} \sinh(\theta\bar{t}) \cos(\phi\bar{t}) + \frac{b - c^2 - c(\theta^2 + \phi^2)}{2\phi(\theta^2 + \phi^2)} \cosh(\theta\bar{t}) \sin(\phi\bar{t})$$

$$a_{24} = \frac{-1}{2\theta} \sinh(\theta\bar{t}) \cos(\phi\bar{t}) - \frac{1}{2\phi} \cosh(\theta\bar{t}) \sin(\phi\bar{t})$$

The state transition matrix has the same structure of that in level 0 for $b = 0$.

Acknowledgments

This research was supported by NASA Langley Research Center under Grant No. NAG-1-939. Daniel D. Moerder served as the grant monitor.

References

[1] Junkins, J. L., "An Asymptotic Perturbation Method for Non-linear Optimal Control Problems," *Journal of Guidance, Control, and Dynamics*, Vol. 9, 1986, pp. 357–367.

[2] 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, 1967, pp. 425–435.

[3] Baldwin, J. F., and Williams, J. H. S., "The Use of a Method of Perturbations in the Synthesis of Closed Loop Optimal Control Laws for Nonlinear Systems," *Automatica*, Vol. 5, 1969, pp. 357–367.

[4] Feeley, T. S., and Speyer, J. L., "A Real-time Approximate Optimal Guidance Law for Flight in a Plane," *American Control Conference Proceedings*, 1990, pp. 2356–2361.

[5] Leung, M. S. K., and Calise, A. J., "An Approach to Optimal Guidance of an Advanced Launch Vehicle Concept," *Proceedings of the American Control Conference*, 1990, pp. 1824–1828.

[6] Leung, S. K., "A Real-Time Near-Optimal Guidance Approach for Launch Vehicles," Ph.D. Dissertation, Georgia Inst. of Technology, Atlanta, GA, March 1992.

[7] Prenter, P. M., *Splines and Variational Methods*, Wiley, New York, 1975.

[8] Hargraves, C. R., and Paris, S. W., "Direct Trajectory Optimization Using Nonlinear Programming and Collocation," *Journal of Guidance, Control, and Dynamics*, Vol. 10, 1987, pp. 338–342.

[9] Leung, M. S. K., and Calise, A. J., "Hybrid Approach to Near-Optimal Launch Vehicle Guidance," *Journal of Guidance, Control, and Dynamics*, Vol. 17, No. 5, 1994, pp. 966–974.

[10] Oberle, H. J., and Grimm, W., "BNDSCO, A Program for Numerical Solution of Optimal Control Problems," English Translation of DFVLR-Mitt. 85–05, 1985.