CHAPTER 7

NUMERICAL SOLUTION OF TWO POINT BOUNDARY VALUE PROBLEMS

7.1 INTRODUCTION

In this chapter, a number of the numerical methods available for the solution of two-point boundary value problems (TPBVPs) in nonlinear ordinary differential equations are presented. Applications of these methods are ilustrated via numerical examples. TPBVPs arising from the application of either classical variational calculus or the Pontryagin's principle to mathematical models of dynamic systems are sometimes *stiff* (in the sense that some particular solutions of the differential equations rapidly increase while others decrease rapidly). Solutions obtained by numerical integration of stiff differential equations often suffer from unacceptably large errors unless sophisticated algorithms designed specifically for these problems are used. Numerical difficulties are not restricted, of course, to stiff systems, however a substantial fraction of the differential equations which govern optimal open loop maneuvers are stiff. A simple stiff system is considered in Example 7.1. The second example considered is an optimal detumble (large angle, nonlinear) maneuver of a rigid asymmetric spacecraft and is non-stiff.

A number of methods for the numerical solution of TPBVPs are currently available, and can be classified under the following categories:

- Quasi-linearization
- Shooting methods
- Polynomial approximation methods
- Finite difference methods
- Hybrid methods
- Continuation/homotopy methods

Of these methods, Quasi-linearization, shooting methods, polynomial approximation, and combination (hybrid) methods are discussed here. An excellent treatment of finite difference methods can be found in Reference 1. The discussion in this chapter is centered around a class of TPBVPs which arises in spacecraft optimal control applications. In Chapter 8, we use these methods (and related methods) to solve for more difficult maneuvers typical of those encountered in practice. More general formulations are presented by Miele (ref. 2) and more recently by Blank and Shinar (ref. 3). These formulations include equality and inequality constraints (both differential and non-differential) on the states, controls, and model parameters of the system.

7.2 STATEMENT OF THE BOUNDARY-VALUE PROBLEM

Let the TPBVP be given by the following sets of nonlinear state ($\mathbf{y}$) and co-state ($\boldsymbol{\lambda}$) nonlinear vector differential equations:

$$\dot{\mathbf{y}} = \mathbf{g}(\mathbf{y},\boldsymbol{\lambda},t), \text{ (n equations)}$$

$$\dot{\boldsymbol{\lambda}} = \mathbf{h}(\mathbf{y},\boldsymbol{\lambda},t), \text{ (n equations)} \qquad (7.1)$$

For a well-posed problem, we need 2n boundary conditions for the solution to Eq. 7.1. We consider split boundary conditions of the form

$$\mathbf{y}(0) = \boldsymbol{\alpha} \text{ (n initial conditions)} \qquad (7.2)$$

and

$$\boldsymbol{\psi}(\mathbf{y}(T), \boldsymbol{\lambda}(T)) = \mathbf{0} \text{ (n final conditions)} \qquad (7.3)$$

where the ψ_i can either be linear or nonlinear functions (constituting a *terminal constraint manifold)*, and $\boldsymbol{\alpha}$ is a known constant vector. Analytical solutions of Eqs. 7.1-7.3 do not exist except for special cases. On the other hand, the split boundary conditions greatly complicate the solution process since direct numerical integration (forward or backward) is not possible; e.g., if $\{\mathbf{y}(0)$ and $\boldsymbol{\lambda}(0)\}$ are specified, then we have an *initial-value problem* in standard form. For the purpose of generalization, we define the augmented 2n vector $\mathbf{x}$, such that

$$\mathbf{x} = [y_1\ y_2\ \cdots\ y_n\ \vdots\ \lambda_1\ \ \lambda_2\ \cdots\ \lambda_n]^T \qquad (7.4)$$

and re-write the TPBVP as

$$\dot{x} = f(x, t) \qquad (7.5)$$

with boundary conditions

$$x_i(0) = \alpha_i \quad , \quad i = 1,2,\ldots,n \qquad (7.6)$$

and

$$\psi(x(T)) = 0 \qquad (7.7)$$

Brief descriptions of various numerical methods for the solution of Eqs. 7.5-7.9 are presented next.

7.3 QUASI-LINEARIZATION

Quasi-linearization is a generalization of the function space Newton-Raphson method in Banach space (ref. 4).

The set of differential equations (Eq. 7.5) are linearized about a nominal solution ($x^k(t)$) for the kth trial solution (i.e., the right-hand side of Eq. 7.5 is expanded in a Taylor's series and only first-order terms are retained). The linearized equations are given by

$$\dot{x}^k(t) + \Delta\dot{x}^k(t) = f(x^k,t) + [\frac{\partial f(x,t)}{\partial x}|_{x^k(t)}]\Delta x^k \qquad (7.8)$$

where $\Delta x^k(t)$ are corrections (departure motion) to the nominal trial trajectory. If $x^k(t)$ is selected such that the initial conditions of Eq. 7.6 are satisfied exactly but the final conditions of Eq. 7.7 are satisfied only approximately, the following boundary conditions are then applicable on $\Delta x^k(t)$:

$$\Delta x_i^k(0) = 0 \quad , \quad i = 1,2, \ldots, n, \qquad (7.9)$$

$$[\frac{\partial\psi}{\partial x}|_{x^k(T)}]\Delta x^k(T) + \psi(x^k(T)) = 0 \qquad (7.10)$$

Since $x^k(t)$ is a known (typically, tabular) function, $\dot{x}^k(t)$ can be approximated directly by numerical differentiation; it is also a simple matter of substitution to recalculate $\dot{x}^k = f(x^k, t)$. Notice that if $x^k(t)$ is obtained by direct integration of Eq. 7.1 with the given initial conditions of Eq. 7.2 and current estimates for the unspecified initial conditions, Eq. 7.8 can be written as

$$\Delta\dot{x}^k(t) = [F^k(t)]\Delta x^k(t) \qquad (7.11)$$

$$[F^k(t)] = [\frac{\partial \mathbf{f}}{\partial x}|_{x^k(t)}] \qquad (7.12)$$

and the boundary conditions of Eqs. 7.9 and 7.10 still hold. This variation avoids numerical differentiation; the differential equations are homogeneous, but at each iteration, Eq. 7.5 must be integrated to obtain $x^k(t)$. The nominal solution for the next iteration is given by

$$x^{k+1}(t) = x^k(t) + \Delta x^k(t) \qquad (7.13)$$

However, to use Eq. 7.11 again, in lieu of Eq. 7.8, we should apply Eq. 7.13 only at the initial time and numerically integrate Eq. 7.5 to obtain a new nominal. A second variation of the method is to substitute Eq. 7.13 in Eq. 7.8 so that

$$\dot{x}^{k+1}(t) = [F^k(t)]x^{k+1} + \{f(x^k(t), t) - [F^k(t)]x^k(t)\} \qquad (7.14)$$

This variation also avoids numerical differentiation, but the nonlinear equations do not have to be integrated at each iteration. The linearized differential equations (Eq. 7.14) are non-homogeneous, and the boundary conditions on $x^{k+1}(t)$ are given by

$$x_i^{k+1}(0) = \alpha_i \quad , \quad i = 1,2,\ldots, n, \qquad (7.15)$$

$$[\frac{\partial \boldsymbol{\psi}}{\partial \mathbf{x}}|_{\mathbf{x}^k(T)}]\mathbf{x}^{k+1}(T) = - \{[\frac{\partial \boldsymbol{\psi}}{\partial \mathbf{x}}|_{\mathbf{x}^k(T)}]\mathbf{x}^k(T) + \boldsymbol{\psi}(\mathbf{x}^k(T))\} \qquad (7.16)$$

Regardless of the variation used, we solve an approximate linear TPBVP at each iteration. For a relatively large class of problems, these approximate solutions converge in the limit to the solution of the nonlinear TPBVP. Roberts and Shipman (ref. 4) present sufficient conditions for convergence of quasi-linearization, estimates of the rate of convergence, and accuracy of the solution, using Kantorovitch's theorem (ref. 5). For most practical purposes, we accept the solutions as converged to a specific tolerance ε if the relative error satisfies

$$\frac{\| x^{k+1}(t) - x^{k}(t) \|}{\| x^{k}(t) \|} \le \varepsilon \tag{7.17}$$

for all integration steps along $x^{k+1}(t)$. The linear TPBVP, at each iteration, can be solved either by shooting methods or polynomial approximation methods.

7.4 SHOOTING METHODS

Shooting methods are iterative methods for adjusting missing initial conditions so that the errors in the terminal boundary conditions in Eq. 7.2 are zero. In this category are method of adjoints, method of complementary functions (ref. 4), methods of particular solutions (ref. 6), and a method using the state transition matrix to compute successive differential corrections (ref. 7). The methods of adjoints and differential corrections are directly applicable to nonlinear problems whereas the other two methods are based on superposition and are valid only for locally linear problems (nonlinear problems are solved by combining these ideas with the quasi-linearization concepts of ref. 8).

7.5 METHOD OF PARTICULAR SOLUTIONS

Consider the linear system of equations:

$$\dot{x} = [F(t)]\ x + D(t) \tag{7.18}$$

with the boundary conditions

$$x_i(0) = \alpha_i \quad , \quad i = 1,2, \ldots, n, \tag{7.19}$$

$$[V]\ x(T) = \beta \tag{7.20}$$

where $[V]$ is a known $[n \times 2n]$ matrix and β is a known constant $n \times 1$ vector. Let

$$x^j = x^j(t) \quad , \quad j = 1,2, \ldots, n+1 \tag{7.21}$$

denote $(n + 1)$ particular solutions obtained by forward numerical solution of

Eq. 7.18 with the following n + 1 sets of initial conditions:

$$x_i^j(0) = \alpha_i \quad , \quad i = 1,2, .., n \quad , \quad j = 1,2, ..., n+1$$

$$x_{n+k}^j(0) = \alpha_k \delta_{jk} \quad , \quad k = 1,2, ..., n \quad , \quad j = 1,2, ..., n+1 \qquad (7.22)$$

where δ_{jk} is the Kronecker delta and α_k is an appropriate trial value (often can taken as unity). Since the differential equation (Eq. 7.18) is linear, we can superimpose the (n + 1) particular solutions to obtain another solution

$$\mathbf{x}(t) = \sum_{j=1}^{n+1} k_j x^j(t) \qquad (7.23)$$

The unknown coefficients (k_j) are obtained to force the solution of Eq. 7.23 to satisfy the bounary conditions of Eq. 7.22. Substituting Eq. 7.23 in the initial conditions of Eq. 7.15, we obtain the side condition:

$$\sum_{j=1}^{n+1} k_j = 1 \qquad (7.24)$$

Similarly, the terminal conditions of Eq. 7.20 yield the n conditions

$$[V] \sum_{j=1}^{n+1} k_j \, x^j(T) = \beta \qquad (7.25)$$

Equations 7.24 and 7.25 constitute n+1 equations which can be solved to determine the (n+1) k_j's. The solution of the linearized TPBVP is then obtained by recombining the individual particular solutions according to Eq. 7.23. Notice that if Eq. 7.11 is used for the linearized TPBVP, due to the nature of the initial conditions in Eq. 7.9, the (n+1)th departure motion is zero. Thus, the (n+1)th particular solution obtained from Eq. 7.13 remains unchanged since Eq. 7.11 is homogeneous. Furthermore, the particular solutions need not be stored; only the state at the initial and final times need be retained. Each iteration involves nonlinear integration of the updated nominal trajectory which will agree exactly with Eq. 7.23 only for the case that the implicit linearization causes negligible errors.

7.6 METHOD OF DIFFERENTIAL CORRECTIONS

This method iteratively refines the initial co-state estimates based on the error in the terminal conditions. It is convenient to use the formulation given by Eqs. 7.1, 7.2, and 7.3 most of the time. The algorithm is summarized in the following steps:

a. The n unspecified initial co-states are selected and Eq. 7.1 is integrated forward in time ($0 \le t \le T$), perhaps (depending upon the particular variation of this method) storing the nominal trajectory $y_i(t)$ and $\hat{\lambda}_i(t)$ along the way. At the final time, $\psi(y(T), \hat{\lambda}(T))$ is computed; if it is zero, it is not necessary to proceed further.

b. Before proceeding to step c, the state transition matrix $\Phi(t,0)$ for the system of Eq. 7.1 has to be computed along the nominal trajectories. The $\Phi(t,0)$ which maps admissible initial condition variations into the first-order predicted variations at time t is given by:

$$[\Phi(t,0)]\Big|_{(\hat{y},\hat{\lambda})} = \begin{bmatrix} \partial y(t)/\partial y(0) & \partial y(t)/\partial \lambda(0) \\ \partial \lambda(t)/\partial y(0) & \partial \lambda(t)/\partial \lambda(0) \end{bmatrix}_{(\hat{y},\hat{\lambda})} \tag{7.26}$$

The evolution of the state transition matrix itself is usually determined by solving the matrix differential equation

$$[\dot{\Phi}(t,0)]\Big|_{(\hat{y},\hat{\lambda})} = [F(t)][\Phi(t,0)]\Big|_{(\hat{y},\hat{\lambda})} \tag{7.27}$$

where F is given by Eq. 7.12 and $\Phi(0,0)$ is a ($2n \times 2n$) identity matrix. If Eq. 7.27 is integrated simultaneously with y, λ of step (a), we can avoid storing $\hat{y}(t)$, $\hat{\lambda}(t)$. We will see that only the upper and lower right portions of $\Phi(T,0)$ are needed in step (c). (The entire Φ matrix is necessary if initial condition continuation (ref. 9) methods are used). $\Phi(t,0)$ can also be approximated by a direct finite difference calculation(ref. 7).

c. The terminal boundary conditions are considered functions of the missing initial co-states. By linearizing about the current estimates of the initial co-states, we obtain the first-order corrections $\Delta\lambda(0)$ to $\hat{\lambda}(0)$ by solving

$$\psi(\hat{y}(T),\ \hat{\lambda}(T)) + \left[\frac{\partial\psi(y(T),\ \lambda(T))}{\partial\lambda(0)}\bigg|_{(\hat{y},\hat{\lambda})}\right]\Delta\lambda(0) = 0 \qquad (7.28)$$

for $\Delta\lambda(0)$, the solution (usually calculated by Gaussian elimination) is indicated formally as

$$\Delta\lambda(0) = -\left[\frac{\partial\psi(y(T),\lambda(T))}{\partial\lambda(0)}\bigg|_{(\hat{y},\hat{\lambda})}\right]^{-1} \psi(\hat{y}(T),\hat{\lambda}(T)) \qquad (7.29)$$

The elements of the partial derivative matrix in Eq. 7.29 can be obtained from $\Phi(T,0)$ and the nominal terminal states and co-states.

d. The previous iterates for the co-state initial conditions are updated by the correction vector $\Delta\lambda(0)$ and the steps (a) through (d) are repeated until convergence of the terminal conditions.

7.7 POLYNOMIAL APPROXIMATION METHODS

Approximation of functions by polynomials has seen widespread use in many engineering and mathematical applications. It has been shown (Weirstrass) that a polynomial of sufficiently high degree can be used to approximate a continuous function to essentially arbitrary accuracy. However, we occasionally encounter subtle obstacles when constructing algorithms to determine polynomial approximations. The discrete points where interpolating polynomials coincide with the function are often chosen at equal intervals in the desired range. However, Fox (ref. 10) gives an example of polynomial interpolation for a simple function $(1 + 25\ x^2)^{-1}$, $-1 \le x \le 1$, for which the equal interval choice leads to serious oscillatory errors between sample points for $|x| \ge .726$, even though the number of sample points and degree of the polynomials are increased without bound. This function, however, can be fit very accurately by linear combination of n Chebyshev polynomials, if the sample points are chosen non-uniformly such that they coincide with the zeros of the

n^{th} Chebyshev polynomial, or, if one uses equally spaced but heavily redundant sample points (and over-determines the polynomial coefficients by the method of least squares). In this section, we are interested in approximating functions by a finite polynomial series, over the entire range and extensions of these ideas to the solution ordinary differential equations. In other applications, functions may be approximated over small subintervals by low order polynomials (e.g., cubic splines), with conditions for continuity imposed at the junctions ("knots") of adjacent sub-intervals in the range. As mentioned previously, quasi-linearization is necessary to apply these ideas to nonlinear TPBVPs.

We denote each polynomial in the series as a basis function. The vector $\mathbf{x}(t)$ in the linear TPBVP of Eqs. 7.18, 7.19, and 7.20 is approximated to the desired accuracy by

$$\mathbf{x}(t) = [\Phi(t)]\mathbf{A} \tag{7.30}$$

where $\Phi(t)$ = Block Diag. $\boldsymbol{\phi}^T(t)$

$$\boldsymbol{\phi}(t) = [\phi_1(t)\phi_2(t)\ \dots\ \phi_{NB}(t)]^T$$

$$\mathbf{A} = [\mathbf{A}_1^T\ \mathbf{A}_2^T\ \dots\ \mathbf{A}_{NB}^T]^T$$

$$\mathbf{A}_i = [a_{i1}\ a_{i2}\ \dots\ a_{iNB}]^T$$

N = 2n, the number of differential equations

and NB is the number of basis functions used. Each of the basis functions $\phi_i(t)$ represents an $(i - 1)^{th}$ degree polynomial in t, forming a linearly independent, complete set. Each of the a_{ij} represents a coefficient such that the i^{th} variable $x_i(t)$ is represented as

$$x_i(t) = \sum_{j=i}^{NB} a_{ij}\ \phi_j(t)\quad,\quad i = 1,2,\ \dots,\ N. \tag{7.31}$$

The $\phi_i(t)$ are specified functions of time (typically, orthogonal polynomials) and their derivatives usually can be obtained through simple recurrence relations. Differentiating Eq. 7.30 we obtain

$$\dot{\mathbf{x}}(t) = \dot{\Phi}(t)\mathbf{A} \tag{7.32}$$

Substituting Eqs. 7.30 and 7.32 into Eq. 7.18, we obtain

$$\dot{\Phi}(t)\mathbf{A} = [F(t)]\Phi(t)\mathbf{A} + \mathbf{D}(t) \tag{7.33}$$

or

$$[\dot{\Phi}(t) - [F(t)]\Phi(t)]\mathbf{A} = D(t) \tag{7.34}$$

which can be written as

$$P_1(t)\mathbf{A} = \mathbf{D}(t) \tag{7.35}$$

with

$$P_1(t) = \dot{\Phi}(t) - [F(t)]\Phi(t),$$

which is an [N x (N x NB)] known time varying matrix. Equation 7.35 can be evaluated at a number (NS) of sample points at times t_i on (T,0) and thereby establish enough equations to determine the coefficient vector **A**. However, we must ensure satisfaction of boundary conditions. Equations 7.19 and 7.20 are repeated as

$$[U]x(0) = \alpha \tag{7.36}$$

$$[V]x(T) = \beta \tag{7.37}$$

where [U] and [V] are known n x N matrices.

Substituting Eq. 7.30 into Eqs. 7.36 and 7.37 yields the two boundary condition equations

$$P_2\mathbf{A} = \alpha \tag{7.38}$$

$$P_3\mathbf{A} = \beta \tag{7.39}$$

where

$$P_2 = [U]\Phi(0) \quad , \quad P_3 = [V]\Phi(T). \tag{7.40}$$

Equation 7.35, evaluated at a number (NS) of sample times (t_i) together with the boundary conditions of Eqs. 7.38 and 7.39 yields a merged linear system

$$P\mathbf{A} = \mathbf{Q} \tag{7.41}$$

where

$$P = \begin{bmatrix} P_1(t_1) \\ P_1(t_2) \\ \vdots \\ P_1(t_{NS}) \\ P_2 \\ P_3 \end{bmatrix}, \quad Q = \begin{Bmatrix} D(t_1) \\ D(t_2) \\ \vdots \\ D(t_{NS}) \\ \alpha \\ \beta \end{Bmatrix}$$

Observe, P is an [N x (NS + 1)] x [N x NB] matrix, whereas Q is an [N x (NS + 1)] x 1 vector. Depending upon the choice of NS, for fixed NB, we can solve Eq. 7.41 either by collocation or the method of least squares (i.e., for NS ≥ NB - 1).

Collocation: NS is selected such that P is a square matrix, i.e., NS = NB - 1. Then the coefficients a_{ij} in **A**, in principle, are determined uniquely by solving the linear system of Eq. 6.41. Thus the solution **x**(t) can be obtained from Eq. 7.30. This solution satisfies the differential equation only at the sample points, however the boundary conditions are satisfied exactly. As NB is increased sufficiently, the average error at intermediate points typically tends to zero.

Least Square Solutions: If NS is selected such that NS > (NB - 1) we can determine the **A** solution using a least squares criterion. This overdetermined solution tends to be smoother and is useful if oscillatory errors are encountered. If the simple least squares method is used, the computed solution generally matches the actual solution at none of the sample points and boundary conditions will not be satisfied exactly, but the square error is typically small uniformly throughout the range. Alternatively, the method of constrained least squares (ref. 11) can be used to enforce exact satisfaction of the boundary conditions. For nonlinear problems, instead of obtaining **A**, we obtain corrections ΔA (the initial a_{ij} are determined easily, since we select the trial nominal profiles $\hat{x}(t)$). The correction vector $\Delta x(t)$ to $\hat{x}(t)$ is expressed as:

$$\Delta x(t) = \Phi(t)\Delta A \tag{7.42}$$

and the updated nominal solution is $x(t) = \Phi(t)\{A + \Delta A\}$. Since the starting solution is typically not very accurate, we may want to restrict the number of basis functions (NB) to a smaller number until the final stages of the solution process. After convergence is achieved for an overly truncated approximation, NB can be increased in an adaptive fashion until convergence to the desired precision is achieved. The linearized TPBVP generally requires more than one correction to converge accurately, for each NB.

The basis functions can be chosen from any linearly independent, complete set of functions, such as Chebyshev or Lengendre polynomials, as well as transcendental basis functions. We consider Chebyshev polynomials in the following examples. There are two popular choices for selecting the sample intervals; uniform intervals are convenient but are usually not optimum from a curve fitting viewpoint, and nonuniform spacing based on the zeros of the basis function polynomials, though better, might be inconvenient for table lookup and interpolation. For Chebyshev polynomials, the zeros occur at higher frequencies near the boundaries (compared to the middle of the normalized time range (-1, 1) over which the polynomials are defined). Hence, such nonuniform interval spacing is better suited to match the true solution in the vicinity of the boundaries.

Another method which exploits orthogonality properties of the basis functions is Galerkin's method, see Meirovitch (ref. 12). But for the present discussion we would need to express elements of [F(t)] and D(t) in Chebyshev series, for example, which would greatly increase the storage requirements.

Recurrence relations for efficient generation and computation of Chebyshev polynomials can be found in Fox (ref. 10).

Quasi-linearization and the method of differential corrections can be thought of as second-order gradient methods. Since both are realizations of a Newton-Raphson method, they converge quadratically but are sensitive to the guessed starting trajectory $\hat{x}(t)$ and the missing initial conditions on $\hat{\lambda}(t)$; i.e., "sufficiently good" estimates are required to converge reliably to the actual solution. The method of particular solutions and polynomial approximation on the other hand, are directly applicable to linear TPBVPs only; these methods can be applied recursively to nonlinear problems, but again, convergence is not guaranteed. The method of polynomial approximation is well suited for many lower-order problems which cannot be solved by methods using numerical integration (e.g., exremely stiff but low-dimensioned systems), for large systems, the drawback is essentially due to the need for the solution of large non-sparse linear systems. For problems which are extremely sensitive to initial conditions, quasi-linearization used in combination with the method of particular solutions has been found better (it converges from poor starting estimates for a wider class of problems) than shooting techniques which iterate using integrations of the parent non-linear equations.

7.8 CONTINUATION

Continuation is a family of numerical methods which augment the capabilities of all the above methods for solving numerically sensitive TPBVPs by decreasing the reliance upon a good initial guess. Discussions and applications can be found in References 4, 9, 16, and 19-24. In essence, a continuous one parameter (α) family of problems is introduced. The family is constructed so that the family degenerates for $\alpha = 0$ to a problem whose solution is available (or can be determined analytically), and for $\alpha = 1$, the family reduces to the problem whose solution is desired. By sweeping α, we can define a large number (if required) of neighboring TPBVPs. By extrapolating from neighboring α converged solutions, the iteration for an intermediate α solution can be initiated with "arbitrarily close" starting estimates for the missing boundary conditions. Unless one encounters a local singular event (e.g., bifurcation points and/or turning points), this approach can very nearly guarantee convergence. The method of neighboring extremals (ref. 7) also belongs to this class of methods since it utilizes the shooting method of differential corrections with gradually imposed boundary conditions. The method of modified quasi-linearization (ref. 13) can be thought of as quasi-linearization in conjunction with continuation.

The "continuous extreme" family of continuation methods (in which *differential equations* are formulated and solved for the missing boundary conditions, as functions of dense α values) is generally referred to as *homotopy* methods. The classical method is due to Davidinko [20] and a recent, very robust algorithm is due to Chow et al. [21-23]. The continuation and homotopy methods clearly trade efficiency (these methods are rather slow, even on problems quickly solved by differential correction methods) for enlarged domains of reliable convergence. These methods are much more "stable with respect to initial ignorance" of the problem's solution. Rather than treat continuation and homotopy methods formally here, we defer to the examples of Chapter 8, 9, and 10 which present several detailed discussions in the context of spacecraft maneuvers.

7.9 MULTIPOINT METHODS

Other methods for stiff and numerically sensitive problems generally solve the TPBVP by converting it into a multipoint boundary value problem; i.e., a finite number (N) of TPBVP's are solved over small sub-intervals $[t_1,t_2]$, ..., $[t_{N-1},T]$ and the solutions are matched at the junctions by continuity conditions. This, in effect, restricts the growth and decay of the solutions over each sub-interval. See for example the works of Ojika et al. (ref. 14), Graney (ref. 15), Deuflhard et al. (ref. 16), Orava and Lautala (ref. 17) and Miele et al. (ref. 18).

7.10 TWO EXAMPLES SOLVED BY THREE METHODS

Example 7.1

A linear TPBVP (ref. 14) is selected to demonstrate the effect of stiffness:

$$\begin{Bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{Bmatrix} = \begin{bmatrix} 1 & 1 \\ k^2 & 1 \end{bmatrix} \begin{Bmatrix} x_1 \\ x_2 \end{Bmatrix} , \; x_1(0) = 1, \; x_2(T) = 0 \tag{7.43}$$

where k is positive and the eigenvalues of the system matrix are 1 ± k. The problem obviously has a stiff behavior when k is large (the solution is governed by two exponentials, one rapidly increasing, the other rapidly decreasing). The analytical solution is:

$$x_1(t) = \frac{1}{2}\{e^{(1+k)t} + e^{(1-k)t}\} - \frac{1}{2}\{e^{(1+k)T} - e^{(1-k)T}\}\{e^{(1+k)t} - e^{(1-k)t}\}\Big/\{e^{(1+k)T} + e^{(1-k)T}\} \tag{7.44}$$

$$x_2(t) = \frac{k}{2}\{e^{(1+k)t} - e^{(1-k)t}\} - \frac{k}{2}\{e^{(1+k)T} - e^{(1-k)T}\}\{e^{(1+k)t} + e^{(1-k)t}\}\Big/\{e^{(1+k)T} + e^{(1-k)T}\} \tag{7.45}$$

It is easy to see that the missing initial condition is $x_2(0) = -5$, for $k = 5$ and $T = 5$. While it is easy to obtain an exact analytical solution, attempting solution of Eq. 7.43 numerically, even for the case of an initial-value problem, will defeat most off-the self differential equation solvers. The analytical solution, the solution obtained by numerical integration (with known initial conditions), and the iterative solution by the method of particular solutions (Section 7.4) are tabulated in Table 7.1. It should be noted that this deceptive "simple" problem is well-known, and most challenging, from the viewpoint of obtaining accurate numerical solutions over long time intervals!

TABLE 7.1

RUNGE-KUTTA AND METHOD OF PARTICULAR SOLUTIONS APPLIED TO EXAMPLE 7.1

TIME	ANALYTICAL SOLUTION		DIRECT INTEGRATION (4 CYC RUNGE-KUTTA, STEP SIZE .001)		METHOD OF PARTICULAR SOLUTIONS (4 CYC RUNGE-KUTTA, STEP SIZE .001)	
t (sec)	x_1	x_2	x_1	x_2	x_1	x_2
0	1.0	-5.0	1.0	-5.0	1.0	-5.0
2.5	.45400E-4	-.22700E-3	.45403E-4	-.22698E-3	.45393E-4	-.22708E-3
5	.41223E-8	0.0	.11329E-1	.56643E-1	-.23438E-1	-.39063E-2

As can be seen from Table 7.1, the direct Runge-Kutta integration and the Method of Particular Solutions produce results which agree with the analytical solution up to errors in the seventh decimal figure (5th significant figure) at 2.5 seconds, but they diverge to no significant digits during the next 2.5 seconds! The missing initial condition is, however, obtained correctly by the method of particular solutions. The Runge-Kutta solution was given the correct initial condition to illustrate the difficulty of forward integration. These errors do not decrease (for 32 bit floating point arithmetic) if the step size is decreased. The Chebyshev polynomial method (Section 6.6) was also applied to this problem. The sample points were selected in two ways, first with equal intervals and secondly, with samples at the zeros of the Chebyshev polynomials.

The Chebyshev polynomial of the first kind is given by

$$T_k(\tau) = \cos(k \cos^{-1}(\tau)) \quad , \quad -1 \le \tau \le 1 \qquad (7.46)$$

where k is the degree of the polynomial and τ is given by

$$\tau = 2t/T - 1 \qquad (7.47)$$

The differentials with respect to τ are related to those with respect to t by

$$\frac{d}{d\tau}(\ ) = \frac{d}{dt}(\ )\frac{dt}{d\tau} = \frac{T}{2}\frac{d}{dt}(\ ) \tag{7.48}$$

The zeros of the Chebyshev polynomial of degree NS-1 are

$$\tau_i = \cos\frac{(i-1)\pi}{NS-1}\ , \quad i = 1, \ldots, NS \tag{7.49}$$

The solution for $x_i(t)$ is obtained as discussed in Section 7.6. The effects of NB and NS on the solution, as well as the effects of uniform versus cosine sampling, are summarized in Table 7.2.

As can be seen, the solution by the method of collocation with cosine sampling is closer to the analytical solution at the boundary though it is slightly in error in the middle of the range (after about 2.5 seconds). With uniform samples, although the numerical solution matches the analytical one closely up to 2.5 seconds and more, however there is a larger error near the final boundary. It can also be seen (from a densely calculated solution) that the solution obtained by the method of least squares has less average error than the collocation solution, especially if we use cosine sampling in the least squares solution.

The convergence of the Chebyshev series can be observed from the rate at which the coefficients a_{ij} decrease. For NS = 29 and NB = 30, the coefficients are given in Table 7.3. These results suggest that polynomial approximation and collocation methods are very advantageous for stiff systems. However, as we mentioned earlier, the large linear system encountered for high dimensioned systems limit this approach to problems of modest dimensionality. Observe, a_{1i} is different from a_{2i} (for tne first few coefficients) by a factor of about -5 for each i (which is the approximate ratio of $x_1(t)$ and $x_2(t)$ up to about 4.0 seconds). Table 7.4 presents the more detailed solutions obtained by collocation and the method of particular solutions, the trajectories $x_1(t)$ and $x_2(t)$ are shown in Figure 7.1. As is evident from Table 7.4, the Chebyshev/collocation numerical solution, for NB = 30, NS = 29, converged accurately to the analytical solution, whereas the method of particular solutions and the Runge-Kutta integrations diverge after t = 3 sec.

TABLE 7.2

CHEBYSHEV POLYNOMIAL SOLUTIONS OF EXAMPLE 7.2

TIME	UNIFORM SAMPLING NS=21, NB=22		COSINE SAMPLING NS=21, NB=22		UNIFORM SAMPLING NS=31, NB=22		COSINE SAMPLING NS=31, NB=22	
(t, sec)	x_1	x_2	x_1	x_2	x_1	x_2	x_1	x_2
0.0	1.0	-5.0	1.0	-5.0	1.0	-5.0	1.0	-5.0
2.5	.45400E-4	-.22700E-3	.45402E-4	-.22701E-3	.45400E-4	-.22700E-3	.45401E-4	-.22701E-3
5.0	3.79 E-7	5.59 E-16	4.09 E-9	2.30 E-16	1.08 E-8	-1.12 E-16	3.96 E-9	-1.48 E-16

TABLE 7.3

CHEBYSHEV COEFFICIENTS FOR EXAMPLE 7.2

i	a_{1i}	a_{2i}
1	.12783	- .63917
2	.24253	-1.21263
3	.20716	-1.03551
4	.15966	- .79830
10	.47506E-2	- .23753E-1
15	.31159E-4	- .15580E-3
20	.48000E-7	- .24000E-6
25	.23274E-10	- .11638E-9
30	.29756E-14	- .20507E-13

TABLE 7.4

TRAJECTORIES FOR EXAMPLE 7.2

	ANALYTICAL SOLUTION (ref. 14)		COLLOCATION NS=29, NB=30 COSINE SAMPLING		METHOD OF PARTICULAR SOLUTIONS	
TIME (sec)	x_1	x_2	x_1	x_2	x_1	x_2
0	1.0	-5.0	1.0	-5.0	1.0	-5.0
.5	.13534E0	-.67668E0	.13534E0	-.67668E0	.13534E0	-.67668E0
1.0	.18316E-1	-.91578E-1	.18316E-1	-.91578E-1	.18316E-1	-.91578E-1
1.5	.24788E-2	-.12394E-1	.24788E-2	-.12394E-1	.24788E-2	-.12394E-1
2.0	.33546E-3	-.16773E-2	.33546E-3	-.16773E-2	.33546E-3	-.16773E-2
2.5	.45400E-4	-.22700E-3	.45400E-4	-.22700E-3	.45394E-4	-.22708E-3
3.0	.61442E-5	-.30721E-4	.61442E-5	-.30721E-4	.60275E-5	-.31888E-4
3.5	.83153E-6	-.41576E-5	.83153E-6	-.41576E-5	-.59605E-5	-.26703E-4
4.0	.11254E-6	-.56265E-6	.11254E-6	-.56265E-6	-.21362E-3	.12207E-3
4.5	.15333E-7	-.75636E-7	.15333E-7	-.75637E-7	-.34180E-2	.21973E-2
5.0	.41223E-8	0.0	.41223E-8	.15113E-15	-.23438E-1	-.39063E-2

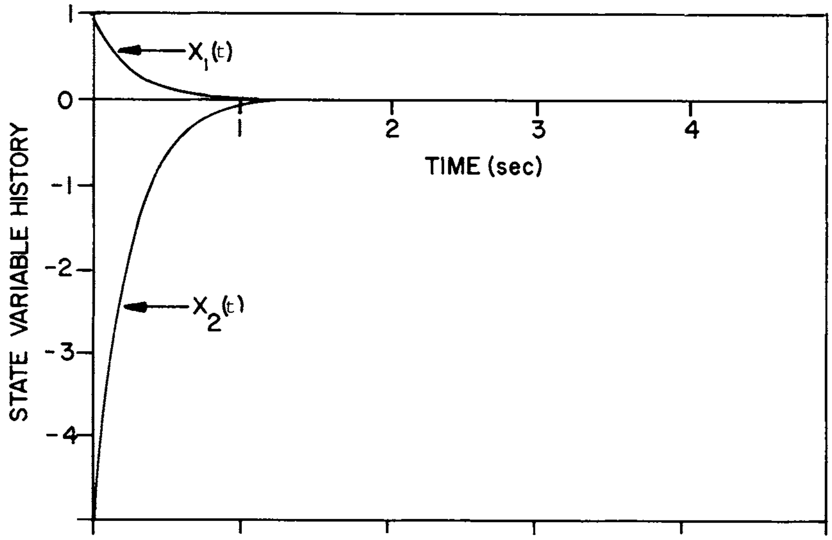

Figure 7.1 Solution Trajectories for Example 7.1

Example 7.2

In this example, we consider optimal detumble maneuvers of a rigid asymmetric spacecraft. The Euler's equations for the angular velocities ω_i of the spacecraft are given by

$$\begin{aligned}\dot{\omega}_1 &= - I_1\omega_2\omega_3 + u_1/I_1 \\ \dot{\omega}_2 &= - I_2\omega_3\omega_1 + u_2/I_2 \qquad (7.50) \\ \dot{\omega}_3 &= - I_3\omega_1\omega_2 + u_3/I_3\end{aligned}$$

where u_i are the control torques, and (I_1, I_2, I_3), the spacecraft principal inertias, and $I_1 = (I_3 - I_2)/I_1$, $I_2 = (I_1 - I_3)/I_2$, $I_3 = (I_2 - I_1)/I_3$, the inertia difference ratios.

A performance index based on control torque histories is selected as

$$J = \frac{1}{2}\int_0^T (u_1^2 + u_2^2 + u_3^2)dt \qquad (7.51)$$

We seek the continuous unbounded controls $u_i(t)$ which minimize J subject to Eq. 7.50 with specified initial and final boundary conditions on the $\omega_i(t)$.

Application of Pontragin's principle (see Chapter 6) leads to a TPBVP constituted by the following six state and co-state differential equations:

$$\begin{aligned}
\dot{\omega}_1 &= - I_1\ \omega_2\omega_3 - \lambda_1/I_1^2 \ , & \dot{\lambda}_1 &= I_2\ \omega_3\lambda_2 + I_3\ \omega_2\lambda_3 \\
\dot{\omega}_2 &= - I_2\ \omega_3\omega_1 - \lambda_2/I_2^2 \ , & \dot{\lambda}_2 &= I_1\ \omega_3\lambda_1 + I_3\ \omega_1\lambda_3 \qquad (7.52) \\
\dot{\omega}_3 &= - I_3\ \omega_1\omega_2 - \lambda_3/I_3^2 \ , & \dot{\lambda}_3 &= I_1\ \omega_2\lambda_1 + I_2\ \omega_1\lambda_2
\end{aligned}$$

We see that we have six nonlinear, gyroscopically coupled differential equations.

Arbitrary initial angular velocity boundary conditions can be prescribed; it is desired that the terminal angular velocities are to be driven to zero. Hence, for the system of six equations, we have six state boundary conditions but no terminal conditions on the co-states are specified. We choose the state terminal boundary conditions

$$\omega_1(0) = 0.01 \text{ r/s} \ , \ \omega_2(0) = 0.005 \text{ r/s}, \ \omega_3(0) = 0.001 \text{ r/s},$$

$$\omega_1(T) = \omega_2(T) = \omega_3(T) = 0, \ T = 100 \text{ sec},$$

and inertias

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

We first apply the polynomial approximation (collocation case) method of Section 7.7, using Chebyshev polynomials as basis functions. Since the initial angular velocities are low and the final time (100 sec) is relatively long, we expect low torque levels. Since the anticipated motions are fairly slow, we also expect a near-linear behavior of the optimal trajectory. One choice for the nominal solution is obtained by integrating Eq. 7.52 with the prescribed initial states and zeros for the initial co-states. On the other hand, it is more convenient for the Chebyshev method (since the coefficients are neartrivial), to use a linear approximation for the states, between the two boundaries; the co-states are again assumed to be zero throughout. Then the Chevyshev coefficients for the linear approximations are

$$\left.\begin{array}{l} a_{i1} = (\omega_1(0) + \omega_i(T))/2 \\ a_{i2} = (-\omega_i(0) + \omega_i(T))/2 \end{array}\right. , \quad i = 1,2,3 \qquad (7.53)$$

the rest of the coefficients are taken as zero.

The Jacobian matrix F required for this nonlinear problem is obtained by differentiation of the Eq. 7.52, leading to

$$F = \begin{bmatrix} 0 & -I_1\omega_3 & -I_1\omega_2 & -1/I_1^2 & 0 & 0 \\ -I_2\omega_3 & 0 & -I_2\omega_1 & 0 & -1/I_2^2 & 0 \\ -I_3\omega_2 & -I_3\omega_1 & 0 & 0 & 0 & -1/I_3^2 \\ 0 & I_3\lambda_3 & I_2\lambda_2 & 0 & I_2\omega_3 & I_3\omega_2 \\ I_3\lambda_3 & 0 & I_1\lambda_1 & I_1\omega_3 & 0 & I_3\omega_1 \\ I_2\lambda_2 & I_1\lambda_1 & 0 & I_1\omega_2 & I_2\omega_1 & 0 \end{bmatrix} \qquad (7.54)$$

The solutions obtained by three different methods agreed to at least seven figures. The missing initial conditions are found to be

$\lambda_1(0) = 0.743733760$

$\lambda_2(0) = 0.361845245$

$\lambda_3(0) = 0.129026881$

Table 7.3 gives the converged coefficients for representing the solution in the form of Eq. 7.30. Table 7.4 compares the numerical values of the analytical solution with the collocation solution and the method of particular solutions. We find that the method of collocation with cosine sample points needs only seven basis functions to represent the solutions to this excellent precision. Figure 7.2a,b shows the optimal trajectories and control torques. The number of iterations and the CPU time on the IBM 370/158 computer are tabulated for each the three methods in Table 7.5.

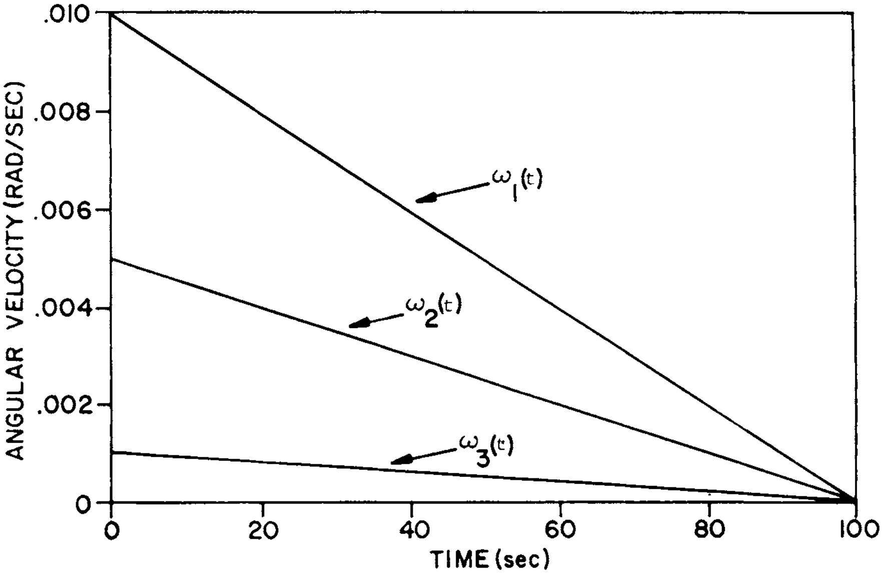

Figure 7.2a Optimal Angular Velocity History for Example 7.2

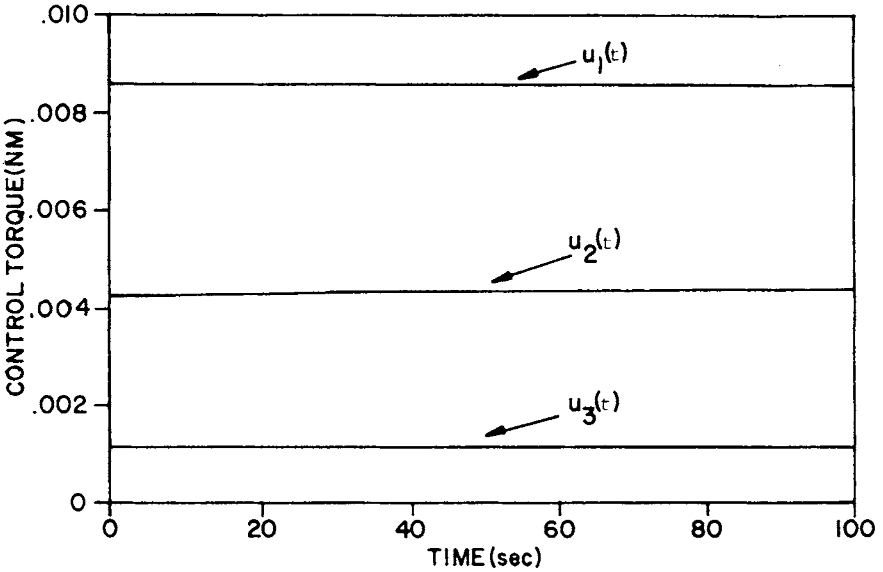

Figure 7.2b Optimal Control Torque History for Example 7.2

The obvious conclusions we can draw from these examples is that the Chebyshev method is certainly superior to the other methods, *for these two examples,* but this conclusion holds only for analogous low-order problems and especially problems whose solutions are accurately representable by a small number of the chosen basis functions. In a general application, lack of prior knowledge of the solutions behavior unfortunately makes broad generalizations very risky. For higher order problems, we find that, as a consequence of computer storage requirements and problems associated with solution of large linear systems, the other two methods prove more widely applicable than the collocation method. We also find [e.g. refs. 16, 19-23] that continuation and homotopy methods vastly enlarge the domain of convergence of the entire family of methods for solving two- and multi-point boundary value problems.

TABLE 7.5

COMPUTATIONAL SUMMARY FOR THREE SOLUTIONS OF EXAMPLE 7.2

	Method of Particular Solutions (Section 6.4)	Differential Correction Shooting Technique (Section 6.5)	Chebyshev Polynomial Collocation (Cosine Sampling) (Section 6.6)
CPU TIME (SEC)	16	11	8
# OF ITERATIONS	5	5	4

7.11 CONCLUDING REMARKS

The present chapter develops formulations and illustrates typical numerical results for three commonly used procedures for solving TPBVPs. These results are generalized and extended for solving problems of higher dimensionality in the subsequent chapters.

We are most indebted to S. R. Vadali for contributing the numerical results of Examples 7.1, 7.2. Vadali's dissertation (ref. 19) provides a number of interesting new applications of these and related methods.

REFERENCES

1. Collatz, L., **The Numerical Treatment of Differential Equations**, 3rd ed. Springer - Verlag, Berlin, 1960.

2. Miele, A., "Recent Advances in Gradient Algorithms for Optimal Control Problems," **Journal of Optimization Theory and Applications**, Vol. 17, Dec. 1975.

3. Blank, D., and Shinar, J., "Efficient Combinations of Numerical Techniques Applied for Aircraft Turning Performance Optimization," **AIAA Journal of Guidance and Control**, Vol. 5, Mar. - Apr. 1982, pp. 124-130.

4. Roberts, S. M., and Shipman, J. S., **Two-Point Boundary Value Problems: Shooting Methods**, American Elsevier Publishing Company, New York, New York 1972.

5. Kantorovitch, L. V., and Akilov, G. P., **Functional Analysis in Normed Spaces**, English Translation, Pergamon prep. 1964.

6. Miele, A., "Method of Particular Solutions for Linear Two-Point Boundary-Value Problems," **Journal of Optimization Theory and Applicaitons:** Vol. 2, No. 4, 1968.

7. Bryson, A. E., Jr., and Ho, Y. C., **Applied Optimal Control**, Halsted Press, Washington, D.C., 1975.

8. Miele, A. and Iyer, R. R., "General Technique for Solving Nonlinear, Two-Point Boundary-Value Problems via the Method of Particular Solutions," **Journal of Optimization Theory and Applications, JOTA:** Vol. 5, No. 5, 1970.

9. Junkins, J. L., and Turner, J. D., "Optimal Continuous Torque Attitude Maneuvers," Paper 78-1400, AIAA/AAS Astrodynamics Conference, Palo Alto, Calif., Aug. 7-9, also AIAA **Journal of Guidance and Control**, Vol. 3, May-June, 1980, pp. 210-217.

10. Fox, L., **Numerical Solution of Ordinary and Partial Differential Equations**, Addison-Wesley, Reading Mass., 1962.

11. Junkins, J. L., **An Introduction to Optimal Estimation of Dynamical Systems**, Sigthoff & Noordhoff, Leyden, The Netherlands. 1978.

12. Meirovitch, L., **Computational Methods in Structural Dynamics**, Sijthoff & Noordhoff, Leyden, The Netherlands, 1982.

13. Miele, A., Mangiavacchi, A. and Aggaruval, A. K., "Modified Quasilinearization Algorithm for Optimal Control Problems with Nondifferential Constraints," **JOTA**, Vol. 14, No. 5, 1974, pp. 529-556.

14. Ojika, T., et al., "A Time-Decomposition Algorithm for a Stiff Linear Two-Point Boundary-Value Problem and Its Application to a Nonlinear Optimal Control Problem," **JOTA** Vol. 27, No. 2, Feb. 79.

15. Graney, L., "The Numerical Solution of Unstable Two-Point Boundary-Value Problems," **Journal of Computational and Applied Mathematics**, Vol. 3, No. 3, 1977.

16. Deuflhard, P., et al., "A Modified Continuation Method for Numerical Solution of Nonlinear Two-Point Boundary Value Problems by Shooting Techniques," **Numerical. Math.**, 26, 327-343 (1976).

17. Orava, P. J., and Lautala, A. J., "Back-and-Fourth Shooting Method for Solving Two-Point Boundary-Value Problems," **JOTA**, Vol. 18, No. 4, April 1976.

18. Miele, A., et al., "Solution of Two-Point Boundary-Value Problems with Jacobian Matrix Characterized by Large Positive Eigenvalues," **Journal of Computational Physics** 15, 117-113 (1974).

19. Vadali, S. R., "Solution of the Two-Point Boundary Value Problems of Optimal Spacecraft Rotational Maneuvers," Ph.D. Dissertation, VPI&SU, May, 1983.

20. Davidenko, D. F., "On a New Method for Numerical Solution of Systems of Nonlinear Equations," **Dokl. Akad. Nauk.**, Vol. 88, pp. 601-602, 1953.

21. Chow, S. N., Mallet-Paret, J., and Yorke, J. A., "Finding Zeros of Maps: Homotopy Methods that are Constructive with Probability One," **Math. Comp.**, Vol. 32, pp. 887-889, 1978.

22. Watson, L. T., and Fenner, D., "Chow-Yorke Algorithm for Fixed Points or Zeros of C^2 Maps," **ACM Trans. Math. Software**, Vol. 6, pp. 252-260, 1980.

23. Dunyak, J. P., Junkins, J. L., and Watson, L. T., "Robust Nonlinear Least Squares Estimation Using the Chow-Yorke Homotopy Method," **AIAA J. of Guid., Control and Dyn.**, Vol. 7, No. 6, pp. 752-755, 1984.