NUMERICAL METHODS

FOR OPTIMAL CONTROL PROBLEMS

Omer Akgiray
BIOSYM Technologies
10065 Barnes Canyon Road
San Diego, California 92121

John Heydweiller
Department of Chemical Engineering & Materials Science
Syracuse University
Syracuse, New York 13244

Prepared for presentation at the 1990 Annual Meeting
of the American Institute of Chemical Engineers,
Chicago

Prepared: September 5, 1990

UNPUBLISHED

1 Introduction

1.1 Problem Description

Many engineering problems require the optimization of differential and algebraic equation models. In particular, these problems appear frequently in reactor design applications. Many of these problems have the following general form:

$$\begin{aligned}
&\underset{Z,U(t)}{\text{Minimize}} \ \phi[Z,U(t),X(t)] = \psi[Z,X(\tau)] + \int_0^\tau h[Z,t,U(t),X(t)]\,dt \\
&\text{subject to :} \\
&\dot{X}(t) = G[Z,t,U(t),X(t)] \ \text{ for } \ t \in [0,\tau] \\
&C[Z,U(t),X(t)] \leq 0 \\
&E[Z,U(t),X(t)] = 0 \\
&X(0) = X_0 \\
&X_L \leq X(t) \leq X_U \\
&U_L \leq U(t) \leq U_U \\
&Z_L \leq Z \leq Z_U
\end{aligned} \qquad \text{(DAOP)}$$

where:
ϕ : objective function
$X(t)$: state vector ($X(t) \in \mathcal{R}^K$)
$U(t)$: control vector
C, E : vectors representing inequality and equality constraints
Z : scalar variables
Due to the presence of both algebraic and differential equations in the model, such a problem is commonly called a *differential-algebraic-optimization-problem* (DAOP). It should be noted that both scalar variables and control functions need to be determined.

1.2 Methods of Solution

In the absence of scalar variables, the DAOP stated above is commonly called an *optimal control problem*. In that case, methods of variational calculus and optimal control theory can be effective. Two commonly applied methods are Pontryagin's maximum principle [18, 26, 27, 36] and dynamic programming [3, 6, 26]. In certain cases, it may also be possible to transform an optimal control problem to a classical variational problem which can then be solved without invoking the maximum principle [35, 36]. Even for an optimal control or variational problem, however, it is often impossible to obtain an analytical solution. A more feasible approach is the numerical solution of the necessary conditions of optimality expressed by the maximum principle. The *steepest descent, variation of extremals,* and *quasilinearization* are three techniques which adapt this approach [26]. The steepest descent method belongs to a class of procedures known as *control vector iteration methods*

[29]. All of these techniques necessiate repeated numerical solutions of two-point boundary value problems, and can therefore be computationally expensive and difficult to converge. Furthermore, these methods cannot easily handle general algebraic constraints on the state functions such as E and C.

In general, scalar variables and algebraic constraints are present in the optimization problem. For such problems, neither nonlinear programming (NLP) methods, nor variational and optimal control methods are applicable directly. This is because NLP methods cannot handle control functions, whereas variational and optimal control methods cannot easily handle scalar variables and algebraic constraints. These problems can only be solved—albeit with a large computational effort—using iterative methods *if* the problem can be decomposed into a number of subproblems each of which is either an NLP problem or an optimal control/variational problem without constraints on the state functions.

2 Global Polynomial Approximation Methods

For the remaining of this paper, attention will be focused on the following *optimal control problem*:

$$
\begin{aligned}
&\underset{u(t)}{\text{Minimize}} \;\; \phi[u(t), X(t)] = \psi[X(\tau)] + \int_0^\tau h[t, u(t), X(t)]\, dt \\
&\text{subject to :} \\
&\dot{X}(t) = G[t, u(t), X(t)] \;\; \text{for} \;\; t \in [0, \tau] \\
&u_L \le u(t) \le u_U \\
&X(0) = X_0
\end{aligned}
\qquad \text{(OCP)}
$$

All the example problems discussed in this paper are of this form. Furthermore, the notation and the presentation of the methods are simplified considerably by using this problem statement. It should be emphasized that most of the approximate solution procedures presented here are easily extended to the more general *differential-algebraic-optimization-problem* given in Section 1.1.

2.1 Rayleigh-Ritz Method

The Rayleigh-Ritz method is a well known approximation method for the solution of variational problems [8]. This method is adapted here by replacing the control function $u(t)$ by an Nth degree polynomial approximation denoted by $u_N(t)$. Writing the polynomial in Lagrangian form, the values of the control function at the interpolation points (t_i) become the parameters in the approximation:

$$
u_N(t) = \sum_{i=1}^{N+1} u(t_i)\ell_i(t) \qquad (1)
$$

$$\text{where} \quad \ell_i(t) = \prod_{k \neq i}^{k=1,N+1} \frac{t - t_k}{t_i - t_k} \tag{2}$$

$$t_i = (i-1)\frac{\tau}{N} \;\; ; \;\; i = 1, \ldots, N+1 \tag{3}$$

With the application of the Rayleigh-Ritz method, an optimal control problem is transformed to a search for the optimal values of the parameters $u(t_i)$ in the approximation. Thus, the OCP given above is transformed to the following nonlinear-programming (NLP) problem:

$$\begin{aligned}
&\underset{u(t_i);i=1,N+1}{\text{Minimize}} \;\; \phi[u_N(t), X(t)] \\
&\text{subject to :} \\
&\dot{X}(t) = G[t, u_N(t), X(t)] \quad \text{for } t \in [0, \tau] \\
&X(0) = X_0 \\
&u_L \leq u(t_i) \leq u_U \qquad ; i = 1, \ldots, N+1
\end{aligned} \tag{NLPP1}$$

This problem can be solved using an NLP algorithm. The solution of the optimal control problem in this transformed form usually is much easier and less expensive than solving the optimality conditions. Furthermore, when the optimal control function is sufficiently smooth, the approximation obtained using the Rayleigh-Ritz method can be made as accurate as the profile obtained by solving the optimality conditions numerically. A modified Rayleigh-Ritz method to handle nonsmooth (e.g. discontinuous) optimal control functions will be presented in Section 3.1.

It should be remarked that the Rayleigh-Ritz method is frequently applied by writing the polynomial approximation in the following form:

$$u_N(t) = a_0 + a_1 t + \ldots + a_N t^N \quad t \in [0, \tau]$$

Another possibility that has been explored is to write the control function approximation in the form of the sums of functions which belong to a complete, orthogonal family of functions, such as trigonometric functions (Bapat,1984)

$$u_N(t) = a_0 + a_1 \cos(s) + \ldots + a_N \cos(Ns) \quad s \in [-\pi, \pi]$$

or Legendre polynomials (Bapat and Heydweiller,1985)

$$u_N(t) = a_0 P_0(s) + a_1 P_1(s) + \ldots + a_N P_N(s) \quad s \in [-1, 1]$$

where $P_i(s)$ is the ith degree Legendre polynomial. [1] With these formulations, the coefficients $a_0, \ldots, a_N$ become the variables in the resulting NLP problem. Their choice of the

[1] To use these formulations, a variable change is affected (from t to s) so that the interval $[0, \tau]$ is mapped onto $[-\pi, \pi]$ or $[-1, 1]$.

locations t_i was also different then the one that is used here (Eqn. 3). For the low degree approximations used ($N \leq 3$), these formulations did not offer a significant improvement in accuracy as compared to the use of the Lagrange polynomial of degree N. Furthermore, since chemical engineering problems involve control functions that represent quantities like catalyst concentration and temperature, using a Lagrange polynomial produces coefficients $u(t_i)$ which are physically meaningful quantities. This becomes useful when initializing the control function approximation, setting bounds on the variables (i.e. $u(t_i)$'s), and interpreting final values of these variables. Another important consideration is that, all $u(t_i)$ are of the same order of magnitute, whereas a_i's generally vary by several orders of magnitute. Therefore, scaling of the variables is better in the resulting optimization problem (NLPP1) when the Lagrangian form is employed.

Returning to the DAOP of Section 1.1, the main incentive for using the Rayleigh-Ritz method is that the search for the optimal values of the scalars $u(t_i)$ can be easily coupled with the search for the optimal values of original scalar variables Z in the problem. Furthermore, when the problem is one of optimizing a chemical process flowsheet, there may be more than one control function to be determined. For example, one may need to determine the optimal temperature profiles for two different chemical reactors within the same process. By replacing each control function by an approximation, the whole problem can be posed as a scalar optimization (NLP) problem. Large scale optimization methods[30] need to be used if this problem has a large number of variables.

2.2 Orthogonal Collocation Method

For a given control profile $u_N(t)$, one has to integrate the differential equations to obtain the state functions which are needed to evaluate the objective function. In general, analytical solutions for these equations cannot be found. A frequent approach, therefore, is the numerical integration [2] of these—possibly nonlinear—ordinary differential equations for each objective function evaluation (i.e. for each new set of values of the variables $u(t_i)$). A large number of objective function evaluations are usually required during optimization calculations carried out to solve NLPP1. Numerical integration may therefore be prohibitively expensive and furthermore, constraints on the state functions cannot be handled easily when this approach is used. Therefore an approximate method, orthogonal collocation [19, 20, 37, 38] is used to solve the differential equations. This method has been shown to give accurate and efficient solutions to differential equations [14, 15, 19, 20]. The method is applied here by writing the state function approximations in the form of the sums of Legendre polynomials:

$$\tilde{X}_M(s) = \sum_{m=0}^{M} A_m P_m(s) \tag{4}$$

[2] By "numerical integration", the use of ODE software employing methods such as Runge-Kutta and multi-step methods is indicated.

$$\text{where} \quad s = \frac{2t - \tau}{\tau} \quad t \in [0, \tau] \tag{5}$$

where:

M : degree of polynomials representing state functions
$P_m(s)$: Legendre polynomial of degree m
A_m : vector of coefficients in the approximation ($A_m \in \mathcal{R}^K$)

The Legendre polynomials

$$P_m(s) = \frac{1}{2^m m!} \frac{d^m}{ds^m}(s^2 - 1)^m \tag{6}$$

are orthogonal on the closed interval $[-1, 1]$ with respect to the weighting function $w(s) = 1$. The first coefficient vector is eliminated here using the initial conditions:

$$\tilde{X}(-1) = \sum_{m=0}^{M} A_m(-1)^m \quad = X_0 \tag{7}$$

Using this equation, the state function approximations can be written in the following form which satisfies the initial conditions:

$$\tilde{X}_M(s) = X_0 + \sum_{m=1}^{M} [(-1)^{m+1} + P_m(s)] A_m \tag{8}$$

The remaining coefficients are calculated by solving the system of simultaneous algebraic equations formed when the residuals are set equal to zero at the roots of the Mth degree Legendre polynomial:

$$\tilde{\Re}(\mathrm{A}, \tilde{u}_N(\sigma_k), \sigma_k) = 0 \quad (k = 1, \ldots, M) \tag{9}$$

where
σ_k : kth root of $P_M(s)$ in [-1,1]
A : $K \times M$ dimensional vector containing A_1 thru A_M

Here $\tilde{\Re}(\mathrm{A}, \tilde{u}_N(s), s) = \Re(\mathrm{A}, u_N(t(s)), t(s))$ is the vector containing the residual functions:

$$\Re(\mathrm{A}, u_N(t), t) \stackrel{\text{def}}{=} \frac{dX_M}{dt} - G[t, u_N(t), X_M(t)] \tag{10}$$

With the application of the collocation method, the differential equations are reduced to a system of algebraic equations (Equation 9). If the original ODE's are linear, this system of algebraic equations will also be linear. In this case, LU factorization [28] is used here to solve for the A_m's. For nonlinear ODE's, these equations can either be solved using the Newton-Raphson method at each objective function evaluation, or they can be included explicitly as equality constraints in the optimization problem. When the equations are solved at each objective function evaluation, the optimization calculations will follow a

feasible path. In the latter case, the coefficients in the state function approximations are included as additional variables in the optimization problem:

$$\begin{aligned}
&\underset{\substack{u(t_i);i=1,N+1\\ A_m;m=1,M}}{\text{Minimize}} \ \phi[u_N(t), X_M(t)] \\
&\text{subject to :} \\
&\tilde{\Re}(\mathrm{A}, \bar{u}_N(\sigma_k), \sigma_k) = 0 \ \ (k = 1, \ldots, M) \\
&u_L \leq u(t_i) \leq u_U \ \ (i = 1, \ldots, N+1)
\end{aligned} \qquad \text{(NLPP2)}$$

This problem is solved here using the successive-quadratic-programming (SQP) algorithm [9]. For nonlinearly constrained problems, the SQP algorithm will in general follow an *infeasible path*, satisfying Eqn. 9 only when the optimization iterations converge to a Kuhn-Tucker point.

2.3 Batch Reactor Optimization Using Global Approximations

The following example illustrates the application of the global approximation methods to a problem with a smooth optimal control function.

Example 1

The example involves a batch chemical reactor in which the following reactions are carried out:

$$A \xrightarrow{k_1} B \xrightarrow{k_2} C$$

Here the objective is to find the temperature policy, for a fixed batch time, which will maximize the production of the intermediate B:

$$\begin{aligned}
&\text{Maximize } c_2(1) \\
&\text{such that :} \\
&\dot{c}_1 = -k_1 c_1^2 \qquad ; \quad c_1(0) = 1 \\
&\dot{c}_2 = k_1 c_1^2 - k_2 c_2 \ ; \quad c_2(0) = 0.5 \\
&T_L \leq T(t) \leq T_U
\end{aligned}$$

where:
c_1 = [A] and c_2 = [B] gmol/L and $k_i = A_{io}\, e^{-E_i/RT(t)}$; i=1,2.

For the constants, the values suggested by Ray [29] were used here (see Appendix).

This problem was studied using various degrees of polynomials for the control and state functions. The exact solution was also obtained by numerically solving the two-point boundary value problem resulting from the optimality conditions for this problem. To assess the effects of the choices of M and N, two sets of calculations were carried out:

1. Having fixed M at a sufficiently large value (what "sufficiently large" means will be discussed next), the best constant, linear (N=1), quadratic (N=2), and cubic (N=3) control functions were determined. These functions are shown and compared to the exact solution on Figures 1 to 4. The objective function values obtained with these profiles are given in Table 1. It may be worthwhile here to add that the accuracies of the state function approximations and objective function values were verified by repeating the calculations using an accurate (but expensive) numerical integration method.

N	objective function
0	1.03267
1	1.03357
2	1.03370
3	1.03372
exact	1.03373

Table 1 : Objective function values for Example 1

2. Having observed that the optimal control function can be approximated by a cubic (N=3) polynomial accurately, another set of optimization calculations were carried out by varying M between 3 and 7, while N=3 was used in each case. The final *control* functions for M=3 and 4 are depicted in Figures 5 and 6. The control function obtained with M=5 is indistinguishable from the profile depicted in Figure 4. It should be noted that in each case a third-degree polynomial was used as the control function, whereas the degree of the polynomials for the state functions was varied. Since increasing M above 5 did not change the results significantly, the corresponding control functions for $M \geq 5$ are not shown here. It can be seen from these results that inaccurate control functions can be obtained if M is too small.

Thus, for this problem, a "sufficiently large" value for M is 5. Increasing M further does not improve the results appreciably, but would increase the computational cost. On the other hand, if M is too small, the errors in the state function approximations lead to incorrect (suboptimal) control functions. This phenomenon is important enough to warrant further discussion. With the application of the orthogonal collocation (or any other approximate) method, the mathematical model of the system (i.e. ODE's) is approximated, and this

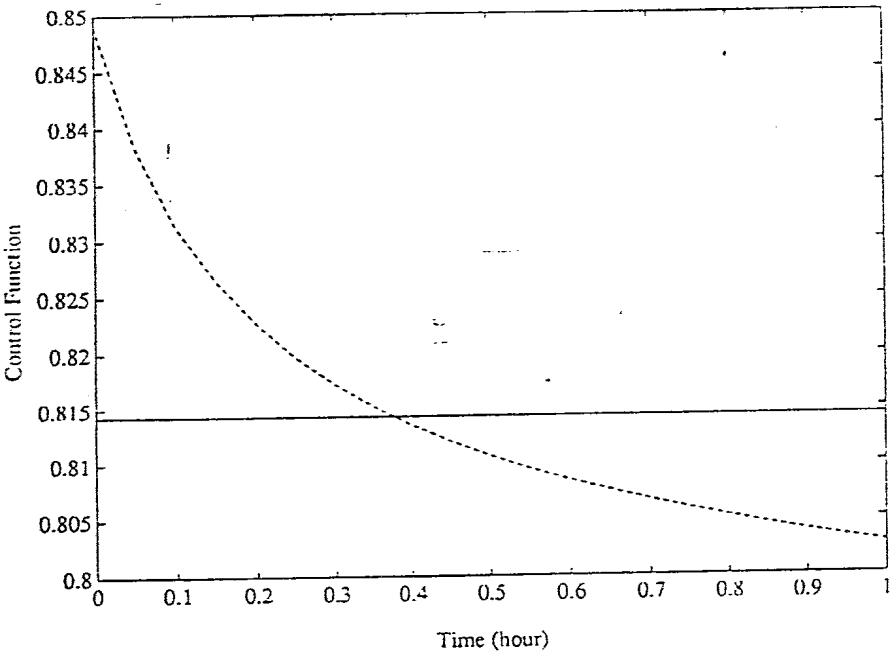


Figure 1: Optimal Constant Temperature (T/T_{max})

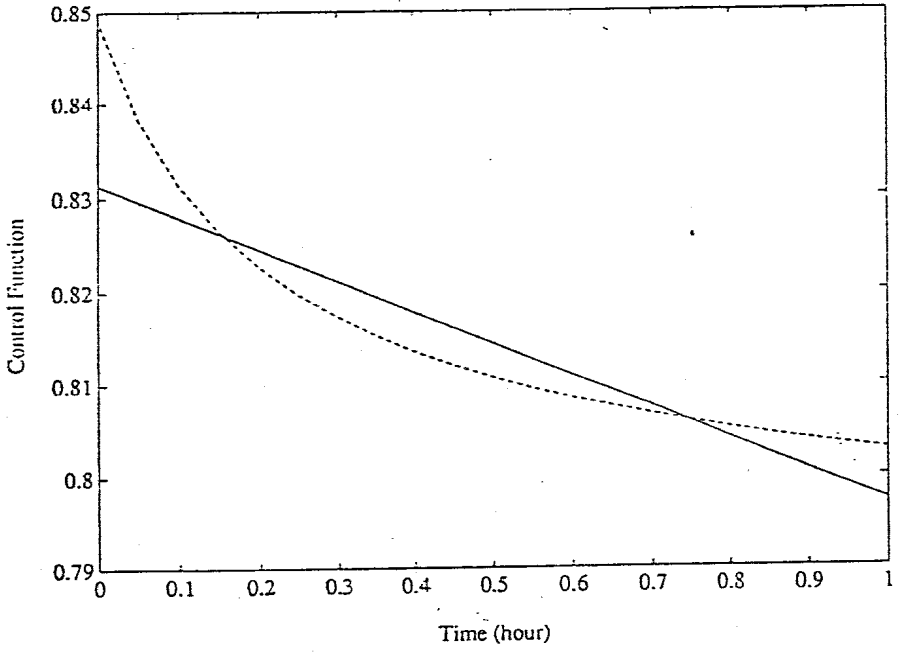


Figure 2: Optimal Linear Approximation ($N = 1$)

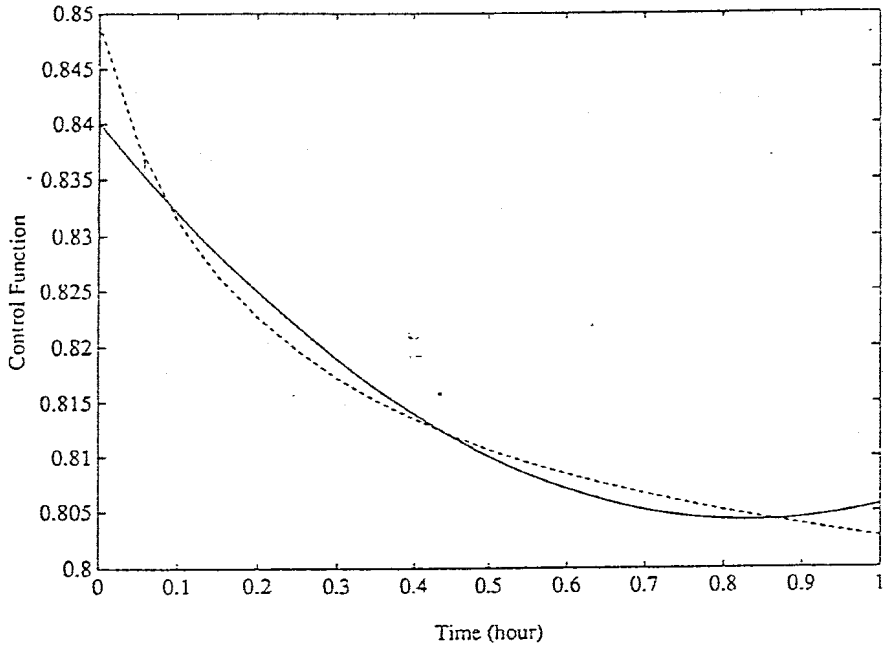


Figure 3: Optimal Quadratic Approximation ($N = 2$)

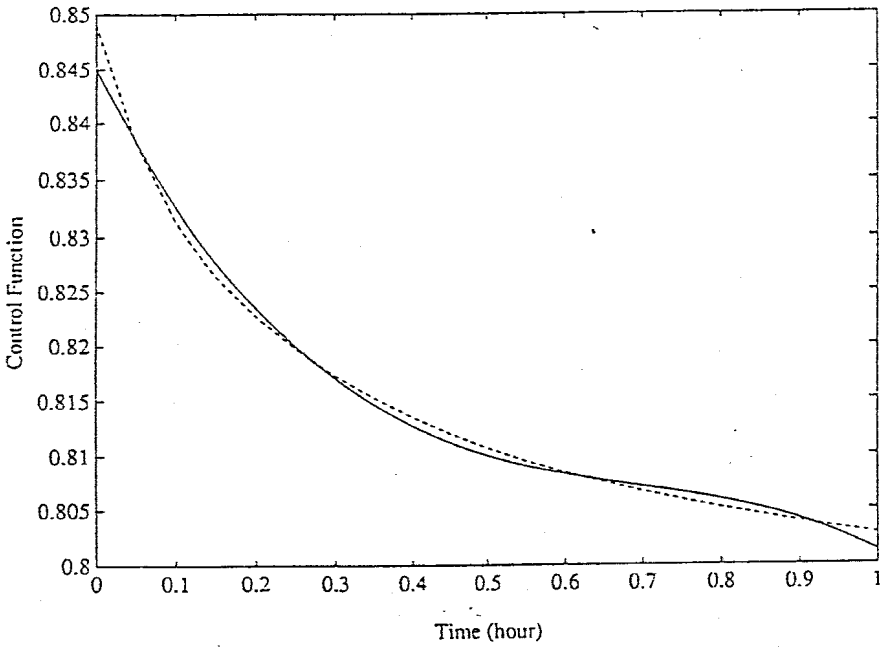


Figure 4: Optimal Cubic Approximation ($N = 3$)

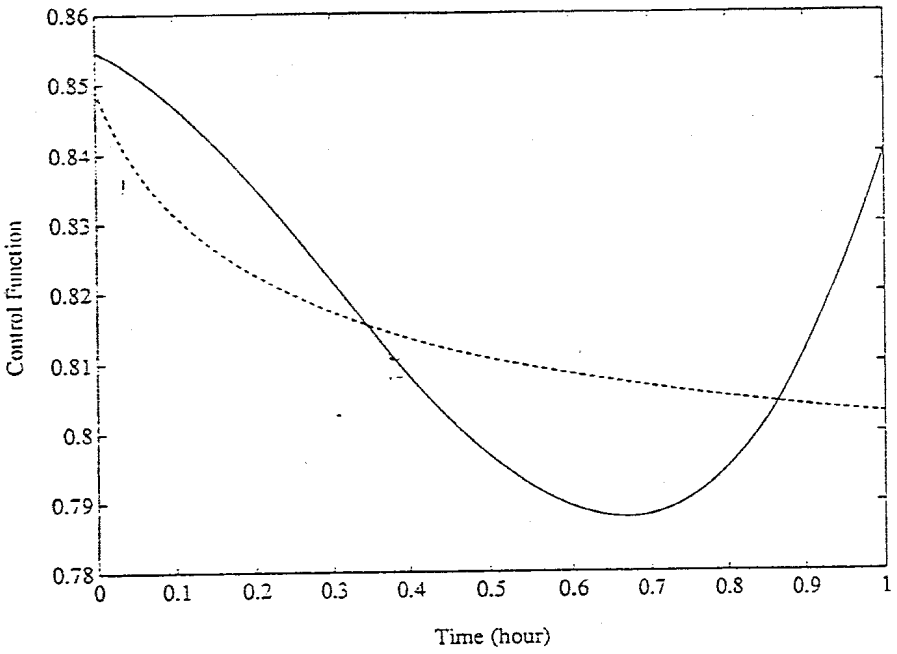


Figure 5: Optimal Cubic Approximation with M=3

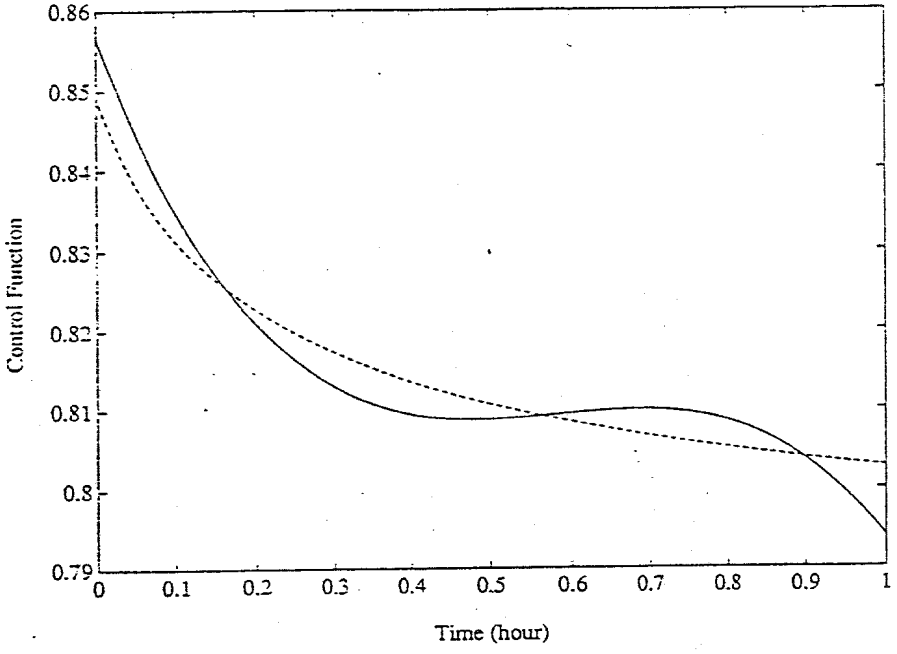


Figure 6: Optimal Cubic Approximation with M=4

approximate model is optimized subsequently. The control function that optimizes this approximate model can be expected to be a good approximation of the optimal control policy for the exact model only if our approximate model is an accurate approximation! That is, only if the state function approximations are accurate. Here the difficulty is twofold: Firstly, one does not know in advance what "accurate" means for a given problem. Put another way, the desired level of accuracy (e.g. the bound on the maximum error in each state function) may depend on the problem being studied. Secondly, the value of M that will yield a given level of accuracy is in general not known. One obvious (and commonly used) approach is to solve the optimization problem repeatedly, increasing M in each case, until no significant change in the results is observed. More efficient strategies such as the use of piecewise polynomials will be considered in the final section of this paper.

2.4 Difficulties in Practice

If the optimal control function is smooth[3], then choosing N and M sufficiently large, the global approximation methods described above will yield accurate approximations of the optimal control function, and the corresponding state functions. This was illustrated with the batch reactor problem. However, for many problems of interest, the optimal control function has corners and/or discontinuities. The approximation of a nonsmooth function by a single polynomial can lead to inaccurate and possibly infeasible results. The following two reactor design problems illustrate these difficulties.

Example 2

The problem is one of determining the blend of two types of solid catalysts along an isothermal tubular reactor to maximize the yield of product C from a chemical reaction of the following form:

$$A \underset{k_2}{\overset{k_1}{\rightleftharpoons}} B \xrightarrow{k_3} C$$

Here one type of catalyst promotes only the first (reversible) reaction, whereas the other type of catalyst promotes only the formation of the product C from B. The mathematical statement of the problem is as follows:

$$\begin{aligned}
&\text{Maximize} \quad x_C(1) = 1 - x_A(1) - x_B(1) \\
&\text{subject to :} \\
&\dot{x}_A = u(t)\,(k_2\,x_B - k_1\,x_A) && ;x_A(0) = 1 \\
&\dot{x}_B = -\dot{x}_A - (1 - u(t))k_3\,x_B && ;x_B(0) = 0 \\
&0 \le u(t) \le 1
\end{aligned}$$

Here x_A, x_B, x_C represent the mole fractions of A, B, and C in the mixture, and k_1, k_2, k_3 are kinetic constants. The control profile $u(t)$ represents the fraction of the catalyst type which

[3]Continuously differentiable on $[0, \tau]$

promotes the first reaction. This problem was first investigated by Gunn and Thomas[21], and was subsequently solved by Jackson[24] using the maximum principle. The following values for the kinetic parameters are used here: $k_1 = 1.0, k_2 = 3.0, k_3 = 0.5$.

Example 3

This problem is similar to the previous batch reactor problem except that the reactions are parallel:

$$A \xrightarrow{k_B} B$$
$$A \xrightarrow{k_C} C$$

Here the objective is to find the temperature policy, for a fixed batch time, which will maximize the production of the intermediate B. For convenience $T(t)/T_{max}$ is used as the control function, instead of $T(t)$. The mathematical statement of the problem is as follows:

$$\begin{aligned}
&\text{Maximize} \quad x_B(1) \\
&\text{subject to :} \\
&\dot{x}_A = -(k_B + k_C)x_A \quad ; x_A(0) = 1 \\
&\dot{x}_B = k_B\, x_A \qquad\qquad ; x_B(0) = 0 \\
&0 \le k_B[T(t)] \le 2
\end{aligned}$$

where k_B and k_C are related to $T(t)$ with the Arrhenius expression (see Appendix).

Both problems were solved using third degree polynomials for the control functions. The final control functions and the exact solutions are shown in Figure 7 and Figure 8. As expected, a control function with a corner or discontinuity cannot be approximated accurately by a single cubic polynomial. It should be added that, for each problem, the same final cubic was obtained when the ODE's were integrated numerically. That is, the inaccuracy of the final control function approximations did not result from errors in the collocation solutions for the state functions. "Incorrect" control functions were obtained because, a single polynomial was used to approximate a nonsmooth control function in each problem.

These and several other problems studied show that the representation of the control function by a single polynomial on $[0, \tau]$ will in general be inadequate when the optimal control function has corners and/or discontinuities. Specifically, a global polynomial approximation for a nonsmooth control function has two major shortcomings:

1. The optimal polynomial is simply a poor approximation of the optimal control function. Consequently, the solution is suboptimal.

2. The optimal polynomial (i.e. the polynomial $u_N(t)$ obtained by solving NLPP1 or NLPP2) may violate the bounds on $u(t)$ at points other than t_i's given by Eqn. 3. This can be seen clearly in Figures 7 and 8. In this case, the solution—as it stands—is of little use, because it is infeasible.

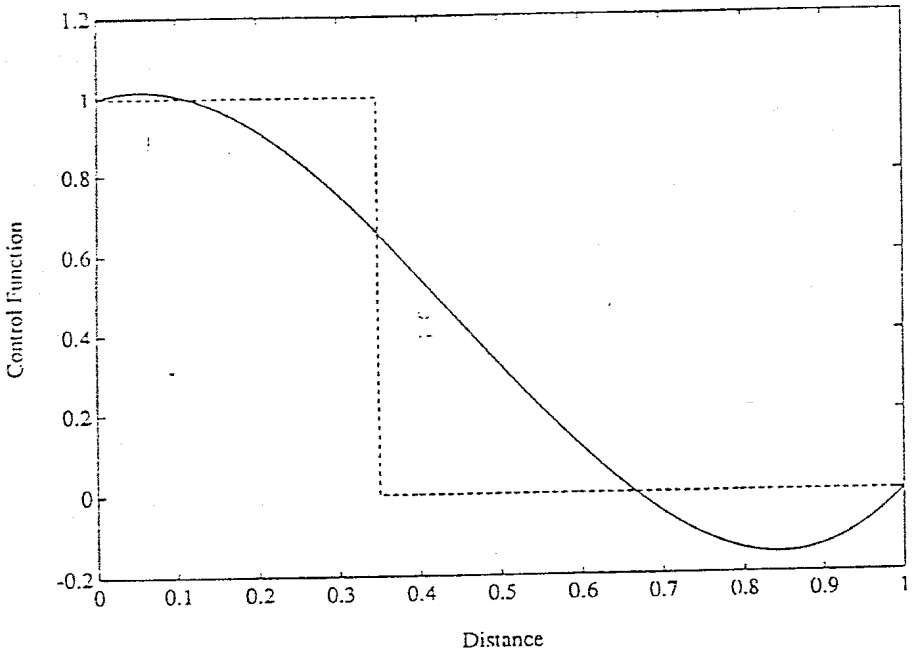


Figure 7: Optimal Cubic Approximation for Example 2

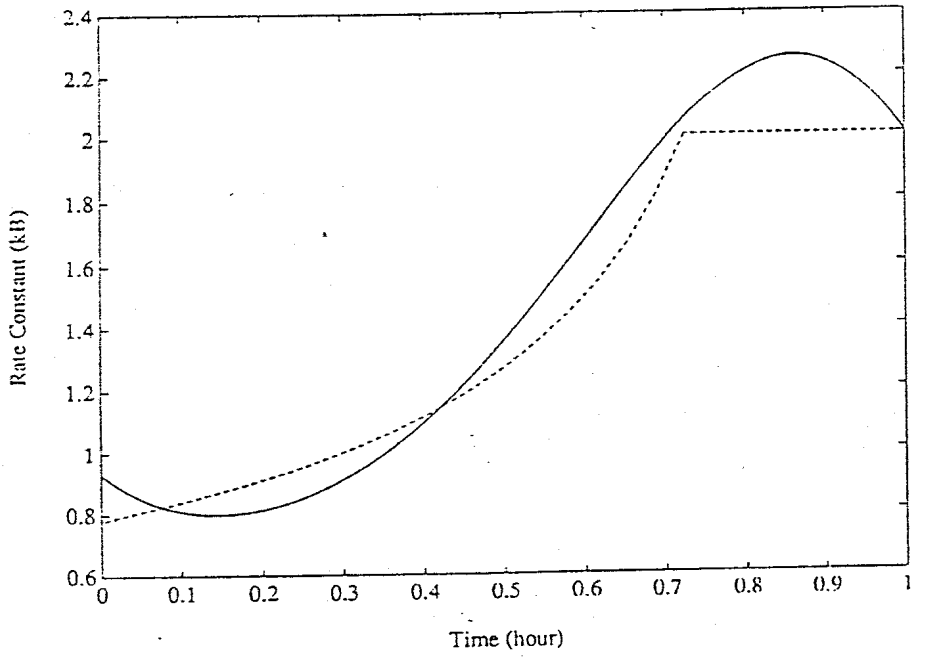


Figure 8: Optimal Cubic Approximation for Example 3

A number of approaches have been taken to overcome the second difficulty, that is, the difficulty in properly bounding the control function approximation:

1. Addition of Inequality Constraints

   One obvious approach would be to impose the bounds not only at the interpolation points, but also at other predetermined locations:

$$u_L \leq u_N(\rho_k) \leq u_u \quad k = 1, \ldots, K_b$$

   where K_b can be chosen as large as desired. It should be noted that $u_N(\rho_k)$ is not one of the variables in the optimization problem (NLPP1 or NLPP2), because ρ_k is not an interpolation point. Therefore, the above inequalities are satisfied by adding $2K_b$ linear inequality constraints to the optimization problem:

$$\sum_{i=1}^{N+1} \ell_i(\rho_k)\, u(t_i) \leq u_U$$

$$-\sum_{i=1}^{N+1} \ell_i(\rho_k)\, u(t_i) \leq -u_L$$

   It is understood that

$$\rho_k \neq t_i \quad i = 1, \ldots, N+1\ ;\ k = 1, \ldots, K_b$$

   for otherwise one or more of the inequalities would be redundant. These inequality constraints, together with the simple bounds on the variables $u(t_i)$, will at least decrease the infeasibility of the polynomial approximation $u_N(t)$ on $[0, \tau]$. Hopefully, with a careful choice of the locations ρ_k, and choosing K_b large enough, $u_N(t)$ will satisfy the bounds everywhere. The shortcoming of this approach is that, while the optimal polynomial approximation may be forced to remain in the admissible region $[u_L, u_U]$, it will still be a suboptimal solution to the original optimal control problem (OCP) when the exact solution has one or more corners or discontinuities. Furthermore, it is usually not obvious *a priori* how K_b and the locations ρ_k should be chosen. To illustrate the application of this method, some of the results obtained for the second example problem (catalyst profile problem) will be given here. When the bounds 0 and 1 on the control function were imposed at the points 0.1, 0.3, 0.5, 0.7 and 0.9, the final control profile (cubic) shown in Figure 9 was obtained. As can be seen in this figure, the bounds are violated to a lesser extent due to the addition of inequality constraints. In a second run, K_b was increased to 15, the points being uniformly distributed between 0.1 and 0.9. Figure 10 shows the final cubic. While the violation of bounds in this case may be negligible, this profile is still a suboptimal solution. Furthermore, as this example illustrates, to make a correct choice of K_b one may have to solve the problem at least twice. If the problem is to be solved for a second time, however, there is a much better approach—namely, the use of piecewise polynomials—which will be outlined later.

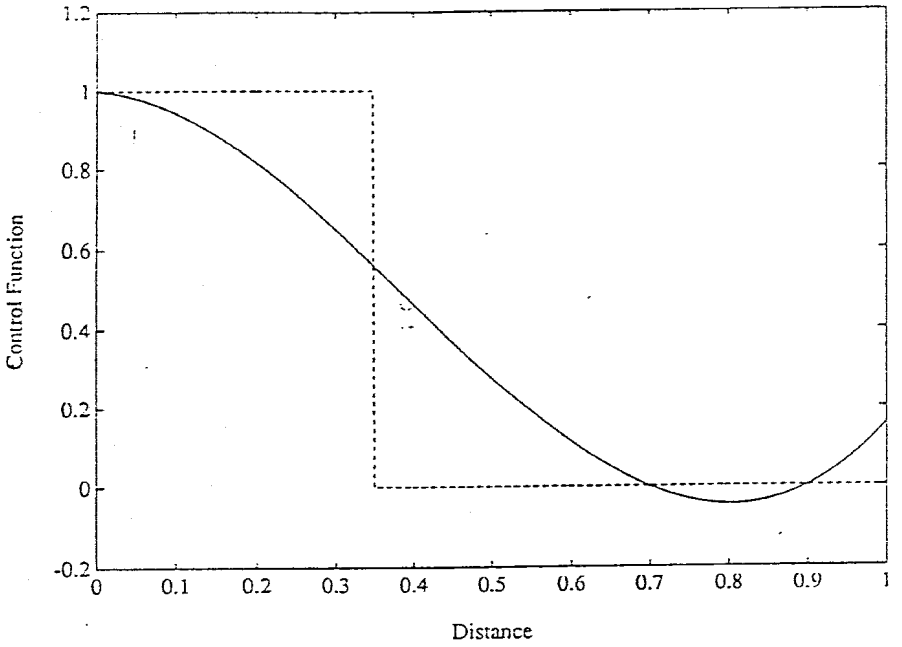


Figure 9: Optimal Cubic Approximation for Example 2 with $K_b = 5$

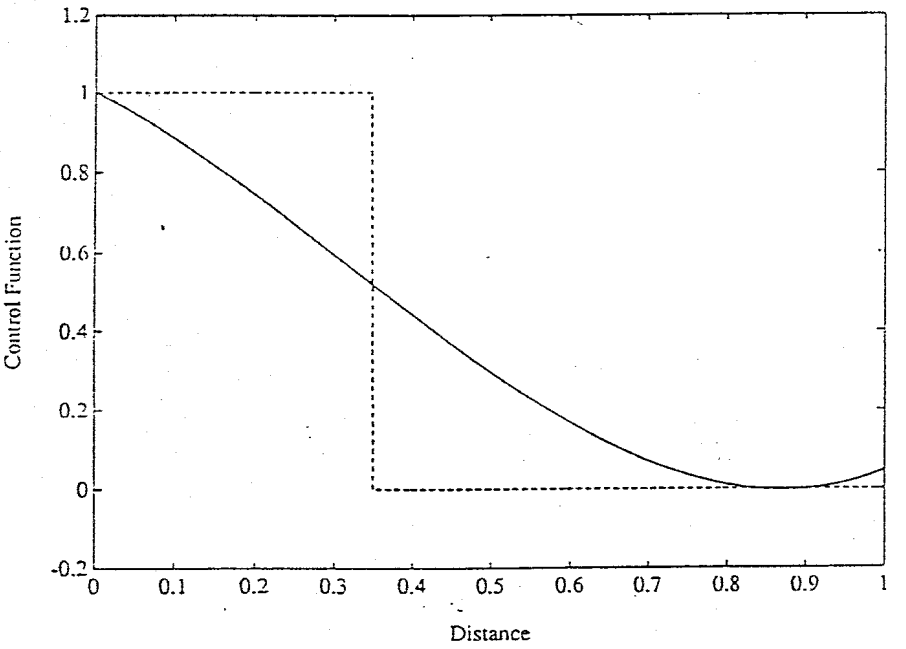


Figure 10: Optimal Cubic Approximation for Example 2 with $K_b = 15$

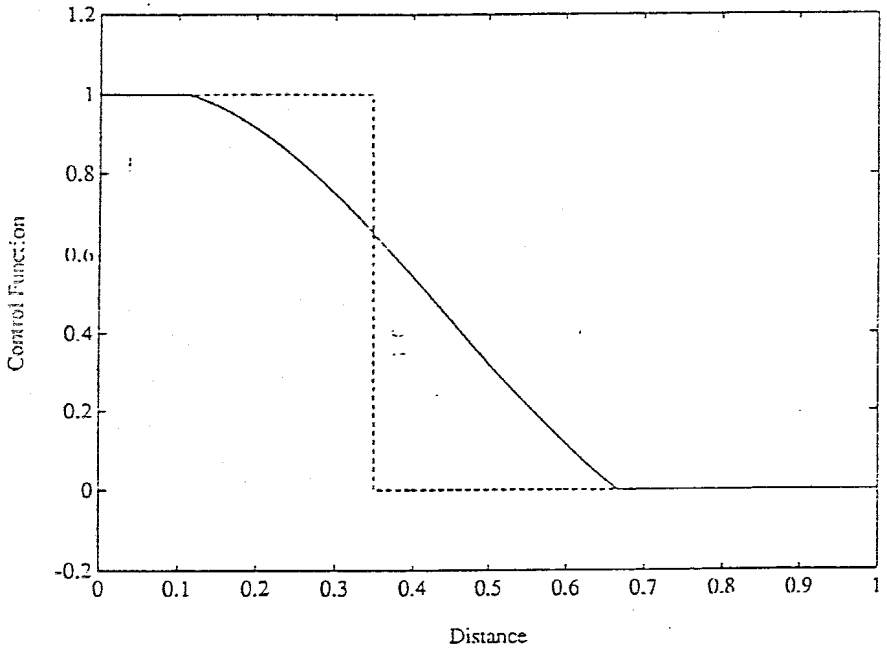


Figure 11: Chopped Cubic Approximation for Example 2

2. Chopping Procedure

In this approach (Cuthrell,1986) one simply *chops* the infeasible sections of the *final* polynomial approximation. The optimization calculations are carried out as before, that is, NLPP1 or NLPP2 is solved to find the best polynomial approximation $u_N(t)$ which is guaranteed to satisfy the bounds only at the interpolation points t_i's. If the final control function violates the bounds, it is modified by chopping its infeasible sections. This "chopped" function is then accepted as the final solution. Finally, the ODE's are integrated numerically to find the state functions, and the value of the objective function corresponding to this final control function. While this approach produces a control policy which satisfies the bounds, it has the same shortcoming as the addition of inequality constraints described above: the final solution will in general be suboptimal. The final profile obtained for Example 2 is shown in Figure 11. This control profile, while feasible, is a poor approximation to the exact solution.

3. Clipping Method

Biegler [10] has studied the batch reactor problem in Example 2 using the following control and state function approximations[4]:

$$u_N(t) = \sum_{i=1}^{N+1} u(t_i)\ell_i(t)$$

[4]His notation is modified here to conform to the notation used in this paper.

$$X_{N+1}(t) = \sum_{i=0}^{N+1} X(t_i) l_i(t)$$

$$l_i(t) = \prod_{k \neq i}^{k=0,N+1} \frac{t - t_k}{t_i - t_k}$$

where

t_0 : initial point ($t_0 = 0$ here).

t_i : (i=1,...,N+1) zeros of the (N+1)-th degree Legendre polynomial defined on $[0, \tau]$. Thus the interpolation points for the control function approximation, and the collocation points are identical in this approach. Having observed that the upper bound u_U is violated between the last collocation (interpolation) point and 1.0, Biegler used a "clipped" control function by simply setting $u(t) = u_U$ in this region. The state functions between the last collocation point and 1.0 were evaluated analytically. An accurate and feasible approximation was obtained in this manner. This method, however, is basically an ad-hoc strategy, and is not generally applicable. For example, the bounds on the control may be violated at points to the left of the last collocation point. The method also requires some a priori information about the optimal control function.

4. The Method of Rosenbrock and Storey

   In [32] Rosenbrock and Storey considered the application of various techniques for the determination of the optimal temperature profile in a tubular reactor. For the particular problem they studied, the optimal temperature profile has a constant segment and a decreasing nonlinear segment. Having obtained such qualitative information about the optimal control function, they used a composite expression in their application of the Rayleigh-Ritz method. This composite expression consisted of a straight line and an exponential with the point of intersection (break point) of the two parts being regarded as an additional variable. They have obtained good results for that particular problem. Their strategy, however, requires a priori knowledge of the solution. Furthermore, they did not considered discontinuous control functions.

While Biegler's Clipping Method, and the method used by Rosenbrock and Storey are basically ad-hoc strategies, and have limited use, these methods are based on certain useful ideas. In Biegler's method, the idea is that the control function should satisfy the bounds everywhere (as opposed to a finite number of points t_i). This is accomplished by "clipping" the infeasible sections. In the latter method, the use of a composite expression for the control function, and the inclusion of the break point as a search variable (to locate the corner) are the important ideas. In the following sections, these ideas will be generalized to design strategies which can be used to solve problems much more general than those considered by these authors.

3 Piecewise Polynomial Approximation Methods

3.1 Modified Rayleigh-Ritz Method

When the optimal control function is smooth, the representation of this function by a single polynomial can be satisfactory. However, if the optimal control function contains corners and/or discontinuities, the Rayleigh-Ritz method with global approximation will fail to produce an accurate control function. This difficulty can be circumvented by dividing the interval $[0, \tau]$ into subintervals $[0, b_1], \ldots, [b_L, \tau]$, and using a different polynomial approximation on each subinterval [1, 2, 5, 16, 22]:

$$u^j_{N_j}(t) = \sum_{i=1}^{N_j+1} u(t_{ij}) \prod_{k \neq i} \frac{t - t_{kj}}{t_{ij} - t_{kj}} \text{ for } t \in [b_{j-1}, b_j] \quad (j = 1, \ldots, L+1) \tag{11}$$

$$t_{ij} = b_{j-1} + (i-1)(b_j - b_{j-1})/N_j \quad (i = 1, \ldots, N_j + 1) \tag{12}$$

where:
N_j : degree of the polynomial approximation on the jth subinterval.
b_j : $(j = 0, \ldots, L+1)$ jth break point. By convention $b_0 = 0$ and $b_{L+1} = \tau$.
t_{ij} : $(i = 1, \ldots, N_j + 1)N_j + 1$ points in the jth subinterval where the values of $u(t)$ are to be determined. In these equations, $u(t_{N_j+1\,j})$ and $u(t_{1\,j+1})$ are interpreted as $u(b_j^-)$ and $u(b_j^+)$, respectively.

The break points (alternatively, the lengths of L subintervals) are made variables in the optimization problem. Furthermore, the control function is allowed to be discontinuous or nondifferentiable at each break point. The following bounds and inequality constraints must also be included in the resulting NLP problem to avoid non-positive subinterval lengths:

$$0 \leq b_j \leq \tau \quad (j = 1, \ldots, L) \tag{13}$$

$$b_j - b_{j+1} + \epsilon \leq 0 \quad (j = 0, \ldots, L) \tag{14}$$

Here ϵ is a small number which prevents the length of the jth subinterval from converging to zero. In some problems, there can be corners but no discontinuities in the control profile. For example, there cannot be a discontinuity in the temperature profile in a single reactor. In this case, continuity must be imposed at the break points:

$$u(b_j^-) = u(b_j^+) \quad (j = 1, \ldots, L) \tag{15}$$

These conditions are satisfied simply by replacing one of the coefficients $u(t_{N_j+1\,j})$ and $u(t_{1\,j+1})$ by the other one in Equation 11.

3.2 Orthogonal Collocation on Subintervals

As can be seen by inspecting the equation $dX/dt = G[t, u, X]$, at a point of discontinuity of the control function $u(t)$, dX/dt is also discontinuous. Furthermore, it is generally expected that $u(t)$, and therefore dX/dt, will change rapidly in a small region around each corner of $u(t)$. Since a single polynomial on $[0, \tau]$ cannot approximate such behaviour of a state function accurately, each state function is approximated by a different polynomial on each subinterval:

$$\tilde{X}_M^j(s_j) = \sum_{m=0}^{M} A_{mj}\, P_m(s_j) \tag{16}$$

$$s_j = \frac{2t - b_{j-1} - b_j}{b_j - b_{j-1}} \text{ for } t \in [b_{j-1}, b_j] \tag{17}$$

where:
s_j : transformed variable for jth subinterval
A_{mj} : vector of coefficients of $P_m(s)$ on subinterval j
$\tilde{X}_M^j$: vector containing the state function approximations on subinterval j

The coefficients are chosen by using continuity of the state functions at each break point, as well as satisfying the initial conditions and collocating the differential equations at M points in each subinterval:

$$\tilde{X}^1(-1) = X_0 \tag{18}$$

$$\tilde{X}^j(1) = \tilde{X}^{j+1}(-1) \quad (j = 1, \ldots, L) \tag{19}$$

$$\tilde{\Re}^j(\mathrm{A}^j, \tilde{u}_N(\sigma_k), \sigma_k) = 0 \quad (k = 1, \ldots, M\ ; j = 1, \ldots, L+1) \tag{20}$$

The expression for the residual function on jth subinterval as a function of s_j is as follows:

$$\tilde{\Re}^j(\mathrm{A}^j, \tilde{u}_N^j(s_j), s_j) = \Big(\frac{2}{b_j - b_{j-1}}\Big) \frac{d\tilde{X}_M^j(s_j)}{ds_j} - \tilde{G}^j[s_j, \tilde{u}_N^j(s_j), \tilde{X}_M^j(s_j)] \tag{21}$$

3.3 Example Problems

Example 2

The optimal catalyst mixing problem was solved using a piecewise linear approximation for the control function. Specifically, the following parameter values were used in Equations 11 and 12:

$$L = 1 \quad \text{(two subintervals)}$$
$$N_1 = N_2 = 1$$

With these, the following were the search variables in the optimization problem:

$$u(0)\ ,\ u(b_1^-)\ ,\ u(b_1^+)\ ,\ u(1)\ ,\ b_1$$

The control function approximation obtained by determining the best values of these parameters is in excellent agreement with the exact solution shown in Figure 7. The results were as follows:

$$u(0) = u(b_1^-) = 1.0 \ , \ u(b_1^+) = u(1) = 0 \ , \ b_1 = 0.35$$

Example 3

For this (second batch reactor) problem, two sets of optimization calculations were carried out:

1. First, the optimal piecewise linear approximation was determined using the following parameters in Equations 11 and 12:

$$\begin{aligned} L &= 1 \quad \text{(two subintervals)} \\ N_1 &= N_2 = 1 \end{aligned}$$

Noting that temperature is a continuous function of time, and therefore

$$u(b_1^-) = u(b_1^+) = u(b_1)$$

the search variables were the following:

$$u(0) \ , \ u(b_1) \ , \ u(1) \ , \ b_1$$

The optimal continuous piecewise linear approximation is shown in Figure 12. The numerical results were as follows:

$$u(0) = 0.9249 \ , \ u(b_1) = u(1) = 1.0 \ , \ b_1 = 0.803$$

The objective function value corresponding to this control policy is 0.5613.

2. Secondly, the optimal continuous piecewise-cubic approximation was calculated using the following values in Equations 11 and 12:

$$\begin{aligned} L &= 1 \quad \text{(two subintervals)} \\ N_1 &= N_2 = 3 \end{aligned}$$

The variables and their optimal values were as follows:

$$\begin{array}{ll} u(0) = 0.9324 & u(\frac{2b_1+1}{3}) = 1.0 \\ u(b_1/3) = 0.9461 & u(\frac{b_1+2}{3}) = 1.0 \\ u(2b_1/3) = 0.9651 & u(1) = 1.0 \\ u(b_1) = 1.0 & b_1 = 0.731 \end{array}$$

These results are in good agreement with the exact solution (Figure 13). The objective function value corresponding to the optimal piecewise-cubic is 0.5619 which is also (for the number of digits shown) the same as the theoretical optimal value.

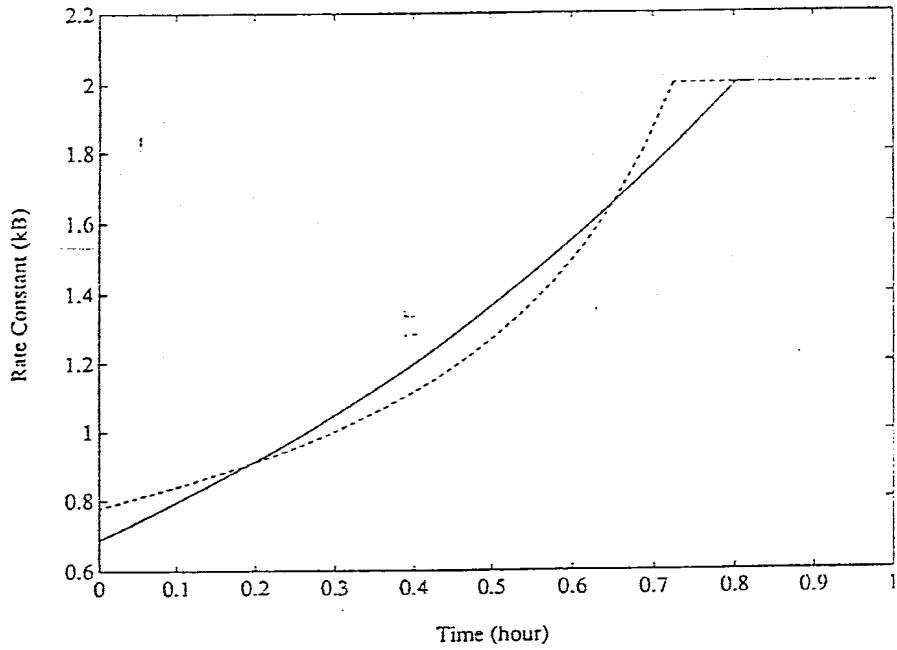


Figure 12: Optimal Piecewise-Linear Control Policy for Example 3

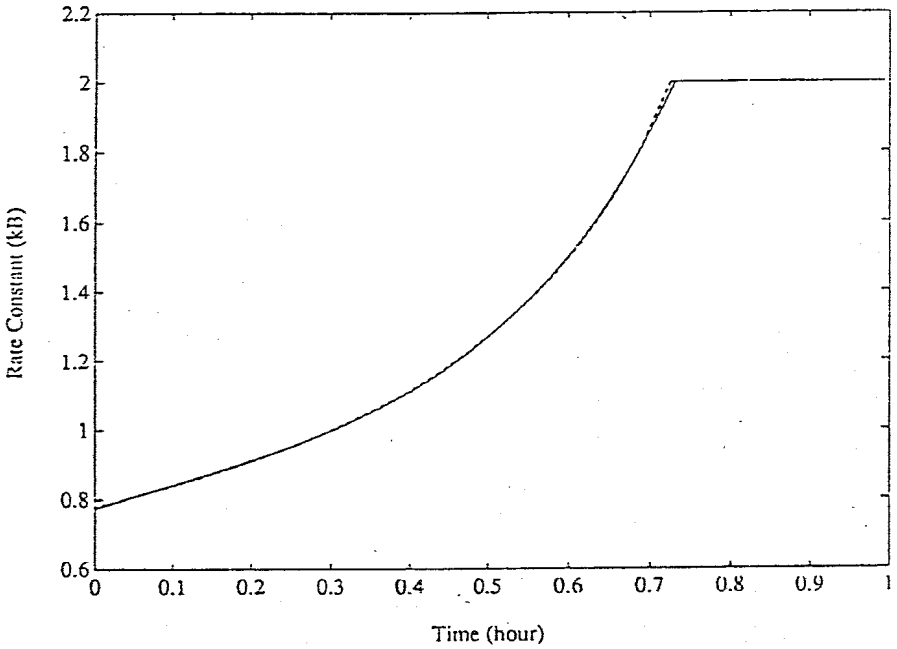


Figure 13: Optimal Piecewise-Cubic Control Policy for Example 3

Example 2′

The optimal catalyst mixing problem (Example 2) was studied so far using the parameter values $k_1 = 1.0, k_2 = 3.0$ and $k_3 = 0.5$. For these particular parameter values, the optimal control function consists of two constant segments and has a single discontinuity. If one or more of these parameters are assigned different numerical values, however, the optimal control policy may assume a noticably different form. For example, when $k_2 = 5$ is used (keeping the other parameters the same), the optimal control function consists of three segments and has two discontinuties. This case, which will be referred to as Example 2′, was studied here to evaluate and illustrate the application of the modified Rayleigh-Ritz method in the presence of multiple discontinuities. The results obtained using two different forms of approximation will be given here: (1) Piecewise-cubic on two subintervals, and (2) piecewise-linear on three subintervals.

1. The problem was first studied using two cubic polynomials on two subintervals. That is, the following parameter values were used in Equations 11 and 12:

$$L = 1 \quad \text{(two subintervals)}$$
$$N_1 = N_2 = 3$$

Depending on the *initial* control approximation (i.e. initial values of the variables b_1 and $u(0), u(b_1/3)$, etc.) used, two different *final* control profiles were obtained for this problem. The values of the variables corresponding to both profiles are shown in Table 2. These are plotted on Figures 14 and 15.

variable	solution # 1	solution #2
b_1	0.4936	0.2516
$u(0)$	1.0	1.0
$u(b_1/3)$	1.0	1.0
$u(2b_1/3)$	0.0	1.0
$u(b_1^-)$	0.0	1.0
$u(b_1^+)$	0.0	0.0
$u(\frac{2b_1+1}{3})$	0.0	0.191
$u(\frac{b_1+2}{3})$	0.0	0.0
$u(1)$	0.0	0.0

Table 2 : Final Values of the Variables for Example 2′
(Piecewise-cubics on two subintervals)

2. A second set of calculations was carried out here to find the optimal piecewise-linear approximation on three subintervals:

$$L = 2 \quad \text{(two breakpoints)}$$
$$N_1 = N_2 = N_3 = 1$$

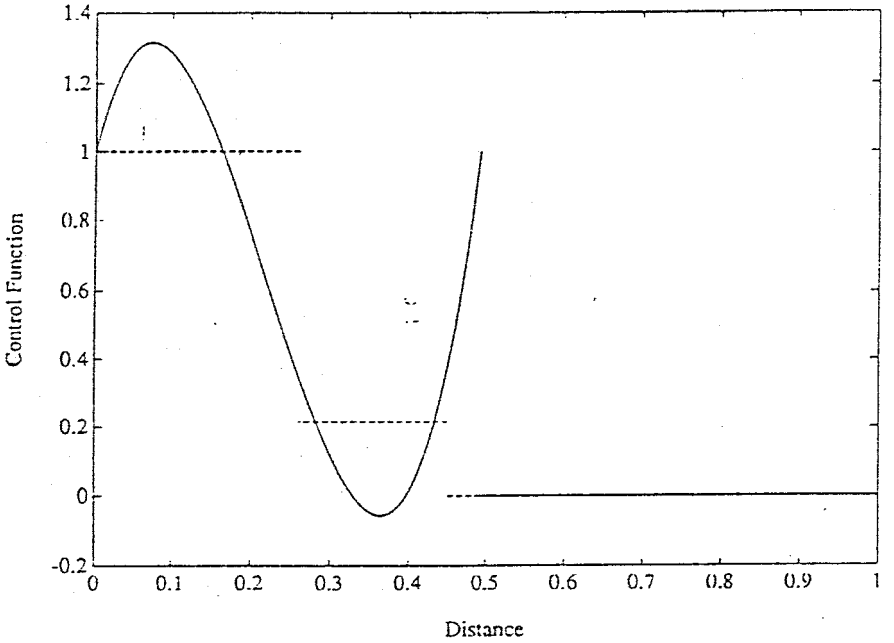


Figure 14: Piecewise-Cubic Solution # 1 for Example 2′

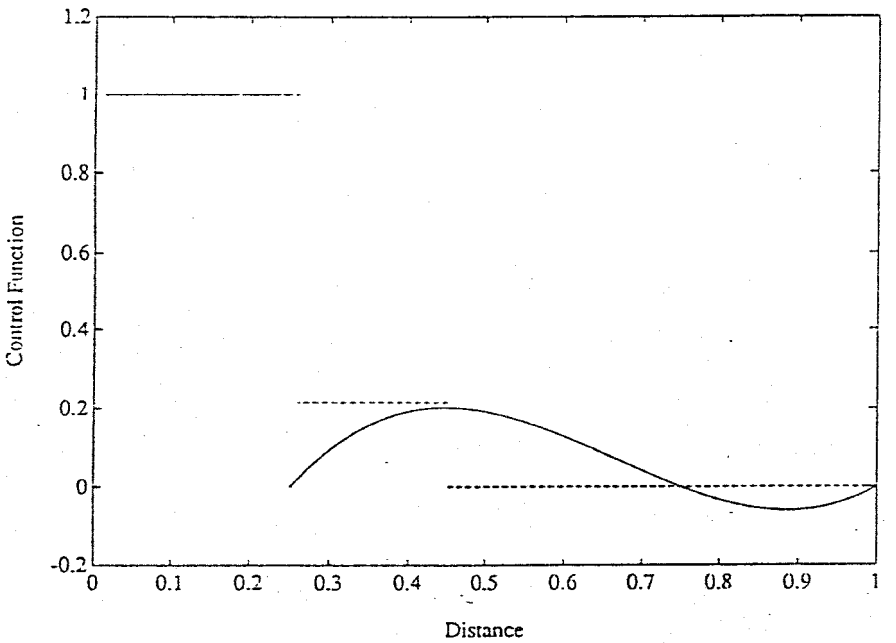


Figure 15: Piecewise-Cubic Solution # 2 for Example 2′

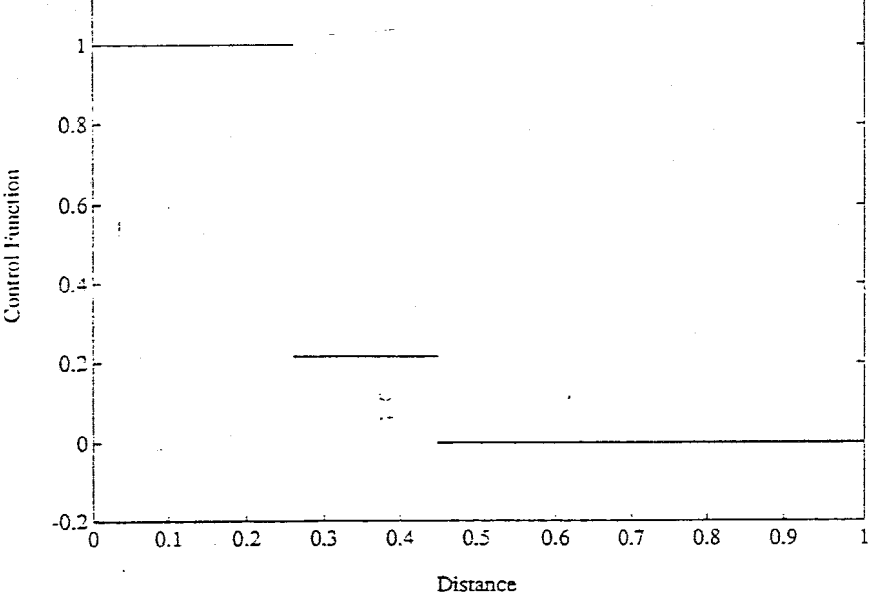


Figure 16: Optimal Piecewise-Linear Solution for Example 2′

The variables and their optimal values are shown in Table 3. The final control approximation is illustrated on Figure 16. This is in excellent agreement with the exact solution.

variable	optimal value
b_1	0.2613
b_2	0.4505
$u(0)$	1.0
$u(b_1^-)$	1.0
$u(b_1^+)$	0.215
$u(b_2^-)$	0.215
$u(b_2^+)$	0.0
$u(1)$	0.0

Table 3 : Final Values of the Variables for Example 2′
(Piecewise-linear approximation on three subintervals)

3.4 Difficulties in Practice

Based on these examples and several other problems studied before [1, 2, 5, 15, 22], the following can be said. The Rayleigh-Ritz method with piecewise approximations can be used to obtain efficient and accurate approximations to non-smooth optimal control functions, *provided that* a sufficiently large number of break points is included in the approximation. In particular, the number of break points L used in the approximation should be greater

than or equal to the total number (say, L') of discontinuities and corners in the optimal control function. Thus, "sufficiently large number of break points" means $L \geq L'$. To assure efficiency, as well as accuracy, the optimal choice would be $L = L'$. This is desirable, because, the number of variables in the resulting NLP problem increases quickly with L. The higher the dimension of an NLP problem, the more difficult and computationally expensive to solve it will be. The main difficulty here is, unfortunately, L' is in general not known a priori. Furthermore, one may not even know whether the optimal control function is smooth, it has one corner/discontinuity, or multiple corners/discontinuities. The next two sections will address this difficulty. In particular, the strategies introduced in Section 4 (Adaptive Approximation Strategy) and Section 5 (Bounded Control Profile Method) have been developed to be used when the presence and/or the number (L') of corners and discontinuities are not known.

4 Adaptive Approximation Strategy

In several problems solved approximately using Eqns. 1 and 3, it has been observed that a cubic polynomial approximation to an optimal control function with a corner or discontinuity develops and retains both a local maximum and a local minimum. Two examples of this behaviour were the "optimal" cubic approximations of the optimal control functions of Examples 2 and 3 (see Figures 7 and 8). Based on this experience, the following heuristic strategy has been adapted [5, 22] for problems in which the control function need not be smooth:

> The control function is approximated by a third degree polynomial using Equations 1 and 3. If the final cubic has two stationary points, it is concluded that a corner or discontinuity in the control function may exist. In this case, the original interval is divided into two subintervals, and a second set of optimization calculations are carried out to determine the optimal piecewise approximation using Equations 11 and 12.

Furthermore, all the problems studied show that these stationary points appear early in the optimization calculations. Consequently, it is possible and more efficient to affect the switch to the piecewise approximation without completing the calculations using the original cubic approximation. It is therefore proposed here that the appearance and retention of *both* a local minimum and a local maximum be used as a criterion for the replacement of a cubic by piecewise polynomials *before* convergence is achieved using the original cubic. The proposed Adaptive Approximation Strategy is given below:

1. Based on all the information available, including physical insight and reasoning, the total number (L) of break points needed is guessed. This guess may often be based on the known solution of a similar problem.

2. In each of the $L + 1$ subintervals, a third degree polynomial approximation ($N = 3$) is used for the control function.

3. The locations of the extrema—if any—of each segment of the control profile are computed after each iteration of the optimization procedure. If *both* a local minimum and a local maximum are observed within a (sub)interval for three consecutive iterations, that (sub)interval is divided into two smaller subintervals, and the original cubic is replaced by piecewise polynomials. The new break point is included as an additional search variable.

4. The search is restarted from a piecewise profile estimated from the third degree polynomial.

Before discussing the results obtained using this procedure, several points need to be made. Firstly, waiting for three consecutive iterations before the switch to the piecewise approximation is arbitrary, and is included in the strategy to assure that the stationary points in the control function approximation are retained. Experience with all the problems studied, however, indicate that once the stationary points appear they are always retained. Secondly, the initial control function approximation should not have both a local maximum and a local minimum (in a subinterval) as this may result in an unnecessary increase in the number of subintervals. Even when the control function was initialized with two stationary points, however, at least one stationary point disappeared quickly in all problems with smooth optimal control functions. Thirdly, it should be emphasized that the Adaptive Approximation Strategy is to be used with Equations 3 and 12 to choose the points $\{t_i\}$ and $\{t_{ij}\}$ in Equations 1 and 11, respectively. When $\{t_i\}$ and $\{t_{ij}\}$ are chosen in a different manner, stationary points may not appear even if the optimal function has a corner or discontinuity. Finally, the original cubic in a subinterval may be replaced by either two cubics or by two linear segments. The choice, which has to be made in advance, generally depends on the following considerations. Piecewise cubic approximation is more accurate when the optimal control profile is nonlinear. It also allows the prediction of another corner or discontinuity within each new subinterval, and the subsequent division of a new subinterval into even smaller subintervals. However, since the criterion for the replacement of the original cubic by piecewise polynomials is likely to be satisfied only when there is a single corner or discontinuity in an original subinterval, this consideration is of less significance. Furthermore, replacement by two cubics increases the dimension of the NLP problem, whereas replacement by two linear segments keeps the dimension nearly constant[5]. Piecewise linear approximation will also satisfy the bounds on the control function (when Eqn.12 is used to choose $\{t_{ij}\}$), whereas a nonlinear approximation may violate the bounds at points other than t_{ij}'s.

The adaptive strategy outlined above was applied to solve Example 2 and Example 3 which were previously solved using global cubic approximations for the control functions. For Example 2, two sets of calculations were carried out. First, using the final cubic approximation shown in Figure 7, a piecewise approximation was estimated and used as the initial profile in the determination of the optimal piecewise linear profiles. That is, the problem was solved a second time using a piecewise approximation after having obtained

[5]The dimension will remain constant if continuity (Eqn. 15) is imposed at the break points.

a converged solution using a global approximation. Second, the Adaptive Approximation Strategy was incorporated into the optimization calculations starting from both of the initial cubic profiles $u = 0.5$ and $u = 1 - t$. The optimization calculations using the global (cubic) approximation were interrupted when the profile had two stationary points in three consecutive iterations. The global approximation at that point was used to estimate an initial profile for the optimization of piecewise linear approximation. The calculations were then pursued to convergence, obtaining the optimal piecewise-linear profile. This final profile (which is the same as that obtained in section 3.3) is in excellent agreement with the maximum principle solution. Table 4 shows the number of SQP iterations performed during each set of calculations.

initial profile	method	with global cubic approx.	with piecewise linear approx.	total
u=0.5	A	17	12	29
	B	3	16	19
u= 1-t	A	8	12	20
	B	3	12	15

Table 4 : Number of SQP iterations for Example 2 (B: Adaptive Approximation)

Similarly, the batch reactor problem in Example 3 was solved with and without the proposed strategy. Savings realized in each case are illustrated in Table 5. It should be noted that the final piecewise cubic profile is in good agreement with the exact solution, and provides an objective function value of 0.5619 which is within 0.001 % of the optimal value. On the other hand, the optimal piecewise linear profile yields an objective value of 0.5613. Furthermore, both piecewise profiles satisfy the bounds, whereas the global cubic approximation violates the upper bound for $0.7 \leq t \leq 1.0$.

initial profile	method	with global cubic approx.	with piecewise linear approx.	total
u= 2	A	19	12	31
	B	3	16	19
$u = 0.25 + 0.75t$	A	18	12	30
	B	8	12	20

initial profile	method	with global cubic approx.	with piecewise cubic approx.	total
u= 2	A	19	19	38
	B	3	24	27
$u = 0.25 + 0.75t$	A	18	19	37
	B	8	20	28

Table 5 : Number of SQP iterations for Example 3 (B: Adaptive Approximation)

5 Bounded Control Profile Method

For a large number of problems with discontinuous and/or nondifferentiable optimal control profiles, one or more sections of the profile are on the boundaries of the admissible control region $[u_L, u_U]$. Using the piecewise approximation method, one has to include a different subinterval for each such section. However, the larger the number of subintervals is, the higher the dimension of the resulting NLP problem will be. The method introduced here keeps the number of subintervals as small as possible, and is proposed as an alternative to the Adaptive Approximation Strategy.

5.1 Description of the Method

The BCP (Bounded Control Profile) Method is essentially another "modified" Rayleigh-Ritz method. Previously, the employment of piecewise polynomials as control function approximations was referred to as the Modified Rayleigh-Ritz Method. The BCP Method improves upon and differs from the Modified Rayleigh-Ritz methods by making the following two important revisions:

1. The method employs the following control function approximation on the jth subinterval:

$$u_b^j(t) = \begin{cases} u_U \text{ if } u_N^j(t) > u_U \\ u_L \text{ if } u_N^j(t) < u_L \\ u_N^j(t) \text{ otherwise} \end{cases} \tag{22}$$

   It will be convenient to use the following compact expression which is equivalent to Equation 22:

$$u_b^j(t) = \text{Max}\{\, u_L, \text{Min}\{u_U, u_N^j(t)\}\,\} \tag{23}$$

2. The bounds $u_L \leq u(t_{ij}) \leq u_U$ are eliminated from the NLP problem[6]. This allows $u_N^j(t)$ to approximate a steep section of the optimal control function accurately.

As will be demonstrated later with an example, the optimal bounded control function obtained with this method is different from the function obtained by chopping the infeasible sections of the optimal polynomial.

It should be noted that the control function approximation used by this method is continuous within each subinterval. However, it is allowed to have upto $2N$ corners within each subinterval *without dividing that subinterval into smaller subintervals*. Furthermore, even though the approximation is restricted to be continuous (within a subinterval), useful results are obtained with the BCP method when there is a discontinuity (in that subinterval). These points will be illustrated shortly using example problems.

There is one potential difficulty in the application of the BCP method when the solutions of the ODEs are approximated using a collocation method. The orthogonal collocation

[6] In computer calculations, the bounds $-10^6 \leq u(t_{ij}) \leq 10^6$ were used.

method constructs the solutions of the differential equations by using the values of the control function at M discrete points on each subinterval. Thus, the shape of the control function at other points than the collocation points does not effect the solutions constructed. For example, if one of the three segments—u_U, u_N^j, and u_L—of the control function occurs on a small section of the subinterval, there may not be any collocation points in that section. In such a case, the state function approximations and the resulting value of the objective function will not be accurate as the effect of that segment will not be taken into account by the collocation method. For the example problems studied here, this difficulty was circumvented by choosing M large enough ($M \geq 5$) to reflect the effect of all segments of the control function approximation u_b^j. So as to generate efficient and accurate solutions of the ODEs, however, the following procedure is used here. At each objective function evaluation (i. e. , for each new set of values of $u(t_{ij}), b_j$) during optimization calculations, the following equations are solved on each subinterval[7]:

$$u_N^j(t) - u_U = 0 \tag{24}$$

$$u_N^j(t) - u_L = 0 \tag{25}$$

Those roots which are in (b_{j-1}, b_j) are the corner points of the control function approximation in subinterval j. The state function approximations are then constructed using orthogonal collocation *on finite elements* within each subinterval. The corner points are used as the end points of these elements. The formulation of collocation on finite elements is exactly the same as that of collocation on subintervals which was outlined before. According to the terminology adapted here, however, a single polynomial is used to construct an approximation to the control function on each subinterval, whereas a distinct polynomial approximates a state function on each finite element. Thus, a subinterval is divided into finite elements to increase the accuracy of the collocation solution when there is at least one corner in the control function within that subinterval.

5.2 Example Problems

Example 3 was first studied using the BCP method, and the following control function (T/T_{max}) approximation was employed:

$$u_b(t) = \text{Min}\{1, \text{Max}\{u_3(t), 0\}\} \tag{26}$$

where $u_3(t)$ is the third-degree polynomial obtained from Equations 1 and 3 using N=3. The search variables were then $u(0), u(1/3), u(2/3)$, and $u(1)$. These variables were left unrestricted, as the bounded function $u_b(t)$ would always satisfy the bounds on the control function whatever the values of the variables. The problem was solved using several different initial profiles (not shown here), and an accurate approximation of the optimal control function was obtained in each case. The final control function obtained using the BCP method is the same as the optimal piecewise-cubic function shown in Figure 13. With

[7]These equations can be solved analytically when $N \leq 3$.

the piecewise-cubic approximation to the control function (that is, with two cubics on two subintervals), the number of variables were 8, that is, twice the number of variables used by the BCP method (with global cubic approximation) to obtain the same final control function. Furthermore, when the BCP method was applied using a global linear approximation (i. e. , Eqn. 26 with $u_1(t)$ instead of $u_3(t)$), the final approximation obtained was the same as the optimal piecewise-linear approximation (Figure 12). The objective function values obtained by various strategies are shown in Table 6.

control profile	objective function
piecewise linear	0.5613
piecewise cubic	0.5619
global cubic with BCP method	0.5619
global linear with BCP method	0.5613
global cubic with chopping	0.5609
exact solution	0.5619

Table 6 : Objective function values for Example 3

The BCP method was also applied to Example 2 (with $k_2 = 3$), using the same form for the control function approximation, i. e. Equation 26. The final approximation obtained is shown in Figure 17. While this profile is not discontinuous in the mathematical sense, it is in excellent agreement with the exact solution except in the small region $[0.34, 0.36]$ near the discontinuity (which is at t=0.35). Furthermore, the objective function value 0.05226 obtained with the BCP method is the same (for the number of digits shown) as the exact optimal value. It should be noted that, the optimal profile obtained using the BCP method clearly indicates a bang-bang solution and a discontinuity at $t \approx 0.35$. Therefore, it is not necessary to solve the problem for a second time using two subintervals. (The control function obtained by chopping the infeasible sections of the final cubic produces an objective function value of 0.05089.)

Finally, the BCP method was applied to Example 2′ ($k_2 = 5$). In this case the optimal control function contains two discontinuities, and consists of three constant segments. The unique feature of this control function is that one of the segments (called the *singular segment*) does not lie on the bounds. Using two cubics and a break point, depending on the initial profile used, one or the other of the control functions in Figures 18 and 19 was obtained. The objective function values that these functions yield are approximately equal to each other, and to the best (exact) value, and are superior to those obtained without the BCP method (see Table 7).

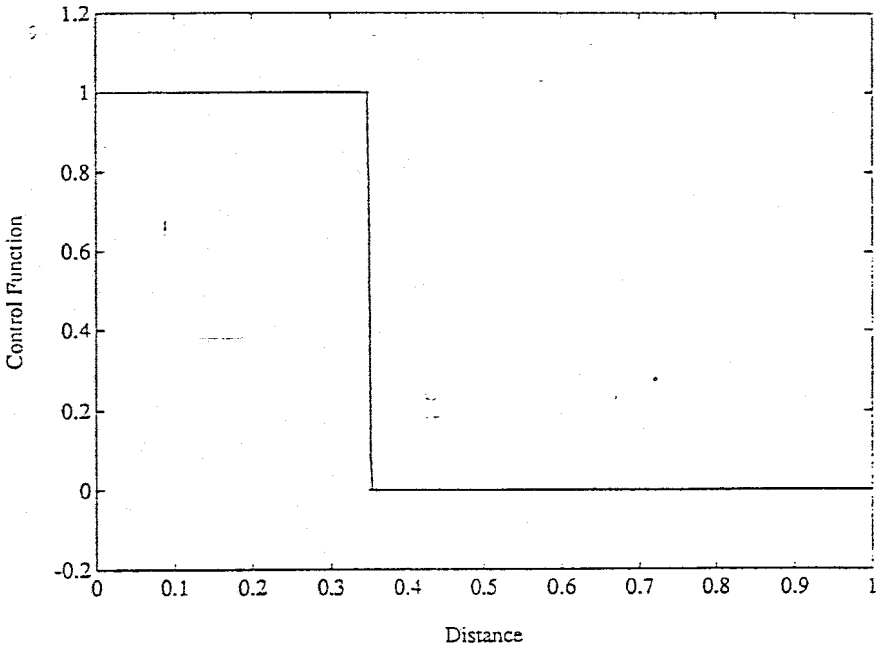


Figure 17: Optimal Bounded Cubic Approximation for Example 2

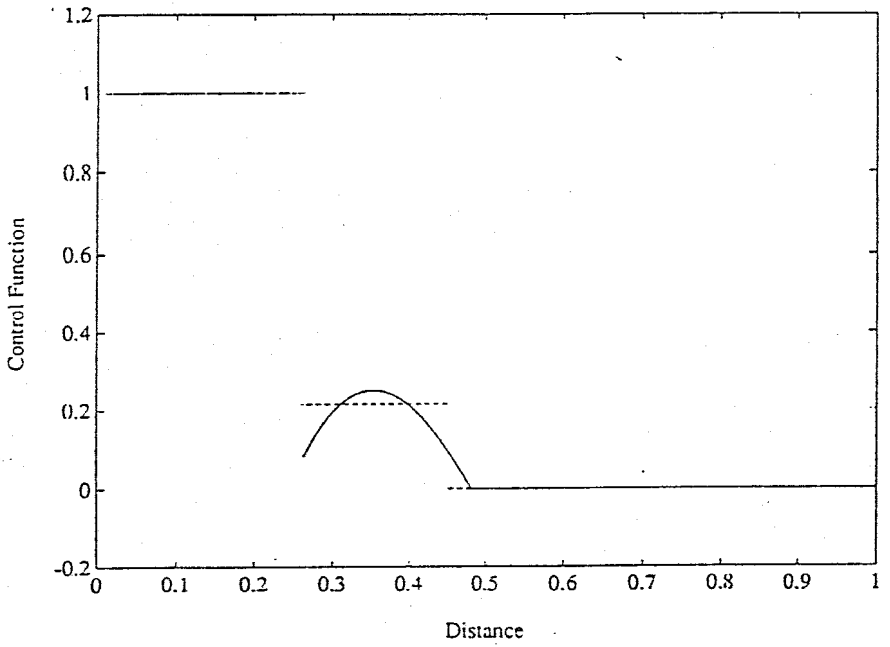


Figure 18: Bounded Piecewise-Cubic Solution # 1 for Example 2′

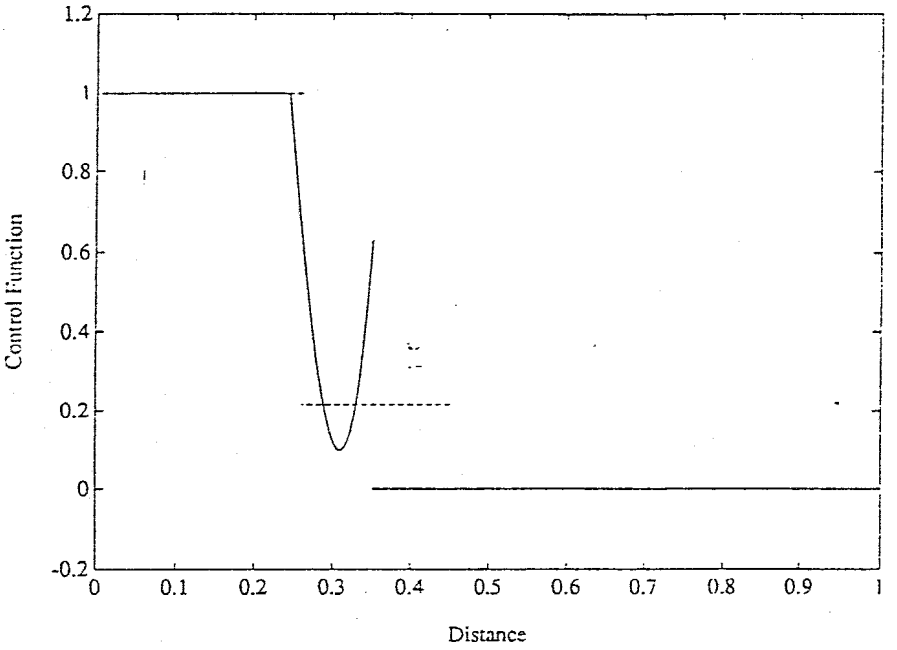


Figure 19: Bounded Piecewise-Cubic Solution # 2 for Example 2′

control profile	Figure	objective function
two cubics with BCP method	Figure 18	0.04113
	Figure 19	0.04113
two cubics with chopping	not shown	0.04091
	not shown	0.04082
three linear segments	Figure 16	0.04113
exact solution	Figures 16,18,19	0.04113

Table 7 : Objective function values for Example 2′ ($k_2 = 5$).

This example illustrates the application of the BCP method with piecewise polynomials. Two cubics were used because the optimal control function for the same problem with a different parameter value ($k_2 = 3$) was found to contain a discontinuity. If the control function were known to consist of three segments with two discontinuities in advance, the best approach would be the use of piecewise polynomials on three subintervals. When such a priori information is lacking, the recommended approach is the use of piecewise polynomials with the BCP method. The number of break points included will be based on the solution of a similar problem, or on some "partial" analytical solution obtained. If the correct number of break points is used, the BCP method will in general not effect the results, and it will not increase computational cost. If too few break points are included, the BCP method yields considerably improved results as compared to any other method.

6 Accuracy and Efficiency Considerations in the Application of the Orthogonal Collocation Method

In section 2.3, the accuracy of the orthogonal collocation method was increased by increasing the degree (M) of the polynomials used for the state function approximations. When a modified Rayleigh-Ritz method was used (section 3.1), orthogonal collocation on *subintervals* was used to avoid "integrating" over corners or discontinuties of the control profile (section 3.2). In this case, again, one can increase M to increase accuracy. Increasing the degree of the polynomials, however, is not necessarily the most efficient approach to obtaining more accurate state functions. An alternative method would be to divide each subinterval into smaller elements, and to use a different polynomial in each element to approximate a state function.[8] This approach is called *orthogonal collocation on finite elements*. While this approach would be more efficient than using a single high-degree polynomial on each subinterval, the basic difficulty remains: In general, one does not know in advance what degree or, how many number of elements is necessary to give sufficiently accurate state function approximations so that a good approximation of the optimal control policy is reached. While this difficulty remains an open question, another question of interest is "what is the strategy that will produce the most accurate state profiles for a fixed level of computational cost?". An approach that is recommended in some recent literature is the use of ortogonal collocation on finite elements with *variable element sizes* [33]. The element sizes are determined such that the error in state profile approximations are (approximately) minimized. Since this approach has been studied elsewhere [15, 16], it will not be detailed here. Another approach (which can be used with variable or fixed element sizes) is the use of *overlapping elements method* of Burka [14]. In this method, the collocation points within an element are distributed such that the last collocation point coincides with the end point of that element. Thus, the residual function will be zero at the end point of each element. The basic motivation for this approach was to avoid the "magnification" of error in passing from one element to the next, because the end point of one element is the initial point of the next one. If each subinterval contains a single element (collocation on subintervals), Burka's method can be expressed mathematically as follows:

$$s_j = \frac{(\sigma_M + 1)t - \sigma_M b_{j-1} - b_j}{b_j - b_{j-1}} \quad \text{for } t \in [b_{j-1}, b_j] \tag{27}$$

This relation, together with Equations 18 and 20, and the following continuity conditions define the orthogonal collocation method on *overlapping subintervals*:

$$\hat{X}^j(\sigma_M) = \hat{X}^{j+1}(-1) \quad (j = 1, \ldots, L) \tag{28}$$

The corresponding expression for the residual function can be obtained from Equation 21 by replacing the factor 2 by $(1 + \sigma_M)$. Similar expressions can be written for the case when

[8] If the BCP method proposed in this paper is used, a subinterval will already be divided into finite elements to avoid integration over the corners within that subinterval. The end points of these elements would be the corner points of the control function. To increase accuracy, these elements may be divided into smaller elements.

each subinterval is further divided into elements, but no new ideas are involved.

All the example problems in this paper were studied using orthogonal collocation on finite elements (or, subintervals) with and without Burka's modification. The results consistently indicated that using overlapping elements decreases the accuracy of the state function approximations, *even at the end points of the elements*. While the residual function vanishes at the end points (and, consequently at $t = \tau$) with the use of overlapping elements, the error in the state functions themselves do not vanish. Furthermore, an important property of the orthogonal collocation method—namely, *superconvergence* at the end points—is not utilized when overlapping elements are used. It is known that, the approximation is of order a^{2M} at the end points of the elements, while the overall approximation is of order a^{M+1} for a first order differential equation [12], where a is the length of the largest element. Thus the approximation is of higher order at the end point as compared to the last collocation point. Therefore, there is also a theoretical reason for using the end point (and, not the last collocation point) as the initial point of the next element.

Returning to the difficulty in choosing the number of elements and/or M, the following may be added. As it was pointed out, an incorrect control function may be obtained due to errors in the state function approximations. On the other hand, choosing M and the number of elements too large would defeat the original purpose of using the collocation method, that is, computational efficiency. For a system of ordinary differential equations such as

$$\frac{dX}{dt} = G[t, X(t)]$$

an obvious strategy is as follows. One can solve the problem once with a certain choice for the number of elements, and solve the problem for a second time using twice the number elements. The difference between these two solutions will in general be a good estimate of the error of the first solution. If this difference is small enough, the second solution can be accepted as the final one. However, for problems of interest here, one has to deal with a system like

$$\frac{dX}{dt} = G[t, u(t), X(t)]$$

where $u(t)$ is not fixed. Therefore the number of elements,etc. needed for sufficient accuracy is also not fixed. That is, as $u(t)$ is changed during optimization iterations, the difficulty of solving the ODE system will also change. Therefore, the need for an accurate *and* efficient solution necessitates an adaptive strategy. That is, the number of elements used in the collocation method sould be varied (adapted) as $u(t)$ changes during optimization calculations. Current work is being done to develop such an adaptive element size strategy which is analogous to the adaptive step size methods used in numerical integration.

APPENDIX

Example 1

$A_{10} = 4000.0$ L/(gmol)(hr)
$A_{20} = 6.2 \times 10^5$/(hr)
$E_1 = 5000$ cal/(gmol)
$E_2 = 10000$ cal/(gmol)
$T_L = 298\ K$
$T_U = 398\ K$
$R = 2$ cal/(gmol)(K)

Example 3

$k_B = 10^6 \exp[-10000/RT(t)]$
$k_C = k_B^2/2$
$x_A = c_A/c_{A_0}$
$x_B = c_B/c_{A_0}$
T : temperature (K)
$c_A(0) = c_{A_0} > 0$
$c_B(0) = c_C(0) = 0$

References

[1] Akgiray, O. and J.C. Heydweiller, *Optimization of Differential Chemical Reactor Models*, presented at the AIChE annual meeting, San Francisco, California (1989).

[2] Akgiray, O. and J.C. Heydweiller, *Adaptive Polynomial Approximations for Optimal Reactor Profiles*, presented at the AIChE annual meeting, Washington D.C. (1988).

[3] Aris, R., *The Optimal Design of Chemical Reactors: A Study in Dynamic Programming*, Academic Press, New York, (1961).

[4] Bapat, S., *Fourier Approximations for Variational Problems in Process Design*, M.S. Thesis, Syracuse University, New York (1984).

[5] Bapat, S. and J.C. Heydweiller, *Adaptive Polynomial Approximations for Process Design*, presented at AIChE annual meeting, Chicago (1985).

[6] Bellman, R., *Dynamic Programming*, Princeton University Press, Princeton, New Jersey (1957).

[7] Bellman, R., *Adaptive Control Processes*, Princeton University Press, Princeton, New Jersey (1961).

[8] Beveridge, G.S.G. and R.S. Schechter, *Optimization: Theory and Practice*, McGraw-Hill, New York (1970).

[9] Biegler, L.T. and J.E. Cuthrell, *Improved Infeasible Path Optimization for Sequential Modular Simulators-II. The Optimization Algorithm*, Comp. Chem. Eng. , 9, 257 (1985).

[10] Biegler, L.T., *Solution of Dynamic Optimization Problems by Successive Quadratic Programming and Orthogonal Collocation*, Comp. Chem. Eng. ,8, 243 (1984).

[11] Vasantharajan, S. and L.T. Biegler, *Simultaneous Strategies for Optimization of Differential-Algebraic Systems*, AIChE Annual Meeting, Washington, DC (1988).

[12] Carl de Boor, *A Practical Guide to Splines*, Springer-Verlag, New York (1987).

[13] Carl de Boor and Blair Swartz, *Collocation at Gaussian Points*, SIAM J. Numer. Analysis, Vol.10, No.4, September (1973).

[14] Burka, M.K., *Solution of Stiff Ordinary Differential Equations by Decomposition and Orthogonal Collocation*, AIChE Journal, 28, 11 (1982).

[15] Cuthrell, J.E. and L.T. Biegler, *On the Optimization of Differential-Algebraic Process Systems*, AIChE Journal, 33, 8 (1987).

[16] Cuthrell, J.E., *On the Optimization of Differential-Algebraic Systems of Equations in Chemical Engineering*, Ph. D. Thesis, Carnegie-Mellon University (1986).

[17] Edgar, T.F. and D.M. Himmelblau, *Optimization of Chemical Processes*, McGraw-Hill, New York (1988).

[18] Fan, L.T., *The Continuous Maximum Principle: A Study of Complex Systems Optimization*, Wiley, New York (1966).

[19] Finlayson, B.A., *The Method of Weighted Residuals and Variational Principles*, Academic Press, New York (1972).

[20] Finlayson, B.A., *Nonlinear Analysis in Chemical Engineering*, McGraw-Hill (1980).

[21] Gunn, D.J. and W.J. Thomas, *Mass Transport and Chemical Reaction in Multifunctional Catalyst Systems*, Chem. Eng. Sci. , 20, 89 (1965).

[22] Heydweiller, J.C. and O. Akgiray, *Adaptive Polynomial Approximations for Optimal Catalyst Profiles*, presented at AIChE meeting, New York (1987).

[23] Holland, C.D. and A.I. Liapis, Computer Methods for Solving Dynamic Separation Problems, McGraw-Hill, New York (1983).

[24] Jackson, R.,*Optimal Use of Mixed Catalysts for Two Successive Chemical Reactions*, J. Opt. Theory Appl. , 2, 27 (1968).

[25] King, R.P., D. Glasser and S.L. Stone, *Optimal Catalyst Concentration Profile for Bifunctional Catalysts*, J. Opt. Theory Appl., 10, 94 (1972).

[26] Kirk, D.E., *Optimal Control Theory*, Prentice-Hall, Englewood Cliffs, New Jersey (1970).

[27] Pontryagin, L.S., V.G. Boltyanskii, R.B., R.B. Gamkrelidze and E.F. Mischenko, *The Mathematical Theory of Optimal Processes*, English translation by K.N. Trigeroff, Interscience, New York (1962).

[28] Press, W.H., B.P. Flannery, Teukolsky, S.A. and W.T. Vetterling, *Numerical Recipes*, Cambridge University Press, Cambridge (1986).

[29] Ray, W.H., *Advanced Process Control*, McGraw-Hill (1981).

[30] Reklaitis, G.V. and I.E. Grossman (editors), *Large Scale Optimization*, Comp. Chem. Eng. , Special Issue, Vol. 7 (1983).

[31] Reklaitis, G.V., A. Ravindran and K.M. Ragsdell, *Engineering Optimization*, John Wiley & Sons, New York (1983).

[32] Rosenbrock, H.H. and C. Storey , *Computational Techniques for Chemical Engineers*, Chapter 10, Pergamon Press, New York (1966).

[33] Russell, R.D. and J.Christiansen, *Adaptive Mesh Selection Strategies for Solving Boundary Value Problems*, SIAM J.Numerical Analysis, Vol. 15, No.1, February (1978).

[34] Schechter, R.S., *The Variational Method in Engineering*, McGraw-Hill, New York (1967).

[35] Troutman, J.L., *Variational Calculus with Elementary Convexity*, Springer-Verlag, New York (1983).

[36] Troutman, J.L., *Optimal Control with Elementary Convexity*, Supplement to [35], Syracuse University, New York (1986).

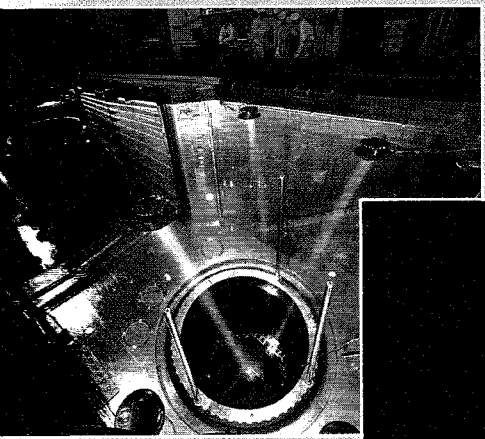
[37] Villadsen, J. and M. Michelsen, *Solution to Differential Equation Models by Polynomial Approximations*, Prentice-Hall, Englewood Cliffs, New Jersey (1978).

[38] Villadsen, J., *Selected Approximation Methods for Chemical Engineering Problems*, Reproset, Copenhagen (1970).

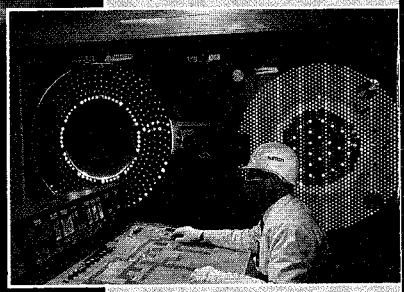
AMERICAN INSTITUTE OF CHEMICAL ENGINEERS

1990 ANNUAL MEETING

November 11-16, 1990
The Palmer House
Chicago, Illinois







GAINING THE COMPETITIVE E D G E

Technical Session Highlights:

- **Thermodynamics**
- **Kinetics, Catalysis & Reactor Engineering**
- **Interfacial Phenomena**
- **Electrochemical Fundamentals**
- **Polymers**
- **Ceramics**
- **Computer Systems Process Design & Control**
- **Food**
- **Biotechnology**
- **Heat Transfer**
- **Education**

For Meeting Registration, Housing & Travel Arrangements, call THE AIChE MEETING HOTLINE 800/225-4570