Simultaneous Optimization and Solution Methods for Batch Reactor Control Profiles

by

J. E. Cuthrell and L T. Biegler

EDRC-6-40-87

SIMULTANEOUS OPTIMIZATION AND SOLUTION METHODS FOR BATCH REACTOR CONTROL PROFILES

by

J. E. Cuthrell[1] and L. T. Biegler

Department of Chemical Engineering
Carnegie-Mellon University
Pittsburgh, Pa 15213

[1]Currently with Shell Development Co. Houston Tx.

Differential-algebraic optimization problems appear frequently in process engineering, especially in process control, reactor design and process identification applications. For fed-batch reactor systems the optimal control problem is especially difficult because of the presence of singular arcs and state variable constraints. For problems of this type we propose a simultaneous optimization and solution strategy based on successive quadratic programming (SQP) and orthogonal collocation on finite elements. In solving the resulting nonlinear programming (NLP) problem, a number of interesting analogs can be drawn to more traditional methods based on variational calculus. First, the collocation method has very desirable stability and accuracy properties. Second, it will be shown that NLP optimality conditions have direct parallels to general variational conditions for optimal control. To demonstrate this strategy, we consider the optimization of a fed-batch penicillin reactor using a number of cases. For the simplest case, the results presented here agree well with previously obtained, analytically-based solutions. In addition, accurate results are presented for more difficult cases where no analytic solution is available.

1. Introduction

The determination of optimal feed rate profiles is an important control problem in many biochemical processes. In many cases, these reactors are operated in fed batch mode and can have slow rates of production or low yield of a high valued product Models of these systems are often nonlinear in the state variables and linear in the feed rate (the control variable). The high valued final product makes determination of an optimal control profile important, especially for product yield maximization. These control profiles are, however, difficult to obtain since problems of this type have optimal control profiles that can be bang/bang and/or have singular arc portions. Additionally, added complexities in the form of constraints on both the control and state profiles are also often present As a result, problems such as these are very difficult, if not impossible, to handle with conventional analytical or numerical solution techniques.

In this paper we present a method which can efficiently solve optimal control problems of this type. In the proposed method differential-algebraic equation (DAE) models are discretized using orthogonal collocation on finite elements with continuous profiles approximated by Lagrange polynomials. The resulting algebraic collocation equations are then written as equality constraints in a Nonlinear Program (NLP) with the polynomial coefficients becoming decision variables. Solution of the NLP results in both determination of the optimal control profile and convergence of the discretized modelling equations. For the purpose of obtaining solutions which have discontinuous control profiles, orthogonal collocation on finite elements is used. Here, by introducing the locations of the finite element knots as decision variables optimal control profiles which are discontinuous can be found.

Solving optimal control problems by discretizing the model to algebraic equations and then solving the resulting optimization problems is not new. For example, Lynn, Parkin and Zahradnik (1970) addressed the simplest optimal control problem (i.e., minimize a function of terminal conditions subject to an ordinary differential equation (ODE) model). Here the necessary conditions were first developed using variational calculus, and then discretized. Later Lynn and Zahradnik (1970) applied the same idea to the problem of optimal feedback control for a distributed system. In this case orthogonal (Chebyshev) polynomials and Galerkin's method were used to discretize a linear partial differential equation (PDE). This resulted in a standard linear-quadratic (LQ) control problem to which the earlier solution method was applied. Oh and Luus (1977) also discretized the variational necessary conditions for both the LQ problem and for a nonlinear model. Here orthogonal collocation was applied at Legendre roots with power series polynomial approximations. Later, Wong and Luus (1982) considered the problem of finding the optimal control profile for a linear parabolic PDE. Here the PDE was discretized using both global orthogonal collocation at Legendre roots and Lagrange interpolation polynomials. The resulting lumped system was then integrated repeatedly with the variables found by direct search optimization. The LQ problem was also considered by Tsang, Himmelblau and Edgar (1975). Here power series polynomial approximations were used and collocation at arbitrarily chosen points (not orthogonal roots) was performed. Solution of the resulting nonlinear program was done with the GRG algorithm. However, all of these studies used a global discretization procedure that is only valid when the optimal state and control profiles are smooth. On problems where discontinuous control profiles are expected, approximate solutions generated with these approaches were poor.

Neuman and Sen (1973a) addressed the more difficult LQ problem with added state variable inequality constraints. This approach uses cubic B-spline basis functions to approximate both the state and control profiles. To provide a more accurate approximation, they use collocation on equally spaced finite elements. The LQ problem is thus reduced to a

Quadratic Program (QP) where the path constraints become just a system of linear constraints enforced at discrete points. The QP solution then yields approximations which are suboptimal, i.e., optimal with respect to the level of approximation. This method was then extended to a distributed parameter system in Neuman and Sen (1973b) where Galerkin*s method was used to discretize the PDE system followed by application of the B-spline approximation to the resulting lumped system.

Finally, to deal with discontinuous control profiles, Sargent and Sullivan (1977) consider a more general optimal control problem that includes inequality constraints dependent upon both the states and control. In this approach the control profile was parameterized over time variable intervals, and the path constraints were transformed into constraints enforced at final time. Instead of converting the ODE model to algebraic equations, a nonlinear program was formulated with control parameters and time intervals, and solved using a gradient based method with state and adjoint equations evaluated by numerical solution of the ODE's for a specified control profile.

Each of the above approaches, while successful at addressing some of the problems associated with optimal control problems, cannot be easily extended to the case considered here. The methods which discretized the problem and then applied a math programming technique often used a relatively poor method of approximation, such as collocation at arbitrarily chosen points or global collocation, and usually did not address the question of approximation accuracy. Conversely, methods that use numerical solution of ODE's can be expensive and are not easily applied to boundary value problems or problems with profile inequality constraints.

In this paper, we present an analysis of the accuracy of our nonlinear programming approach and show why this approach can handle discontinuous control profiles correctly. Theoretical justification of this method is shown through equivalence of the Kuhn-Tucker conditions of the NLP to discretized variational optimality conditions. In addition stability and accuracy properties of orthogonal collocation are also discussed. This results in an equivalence between orthogonal collocation and solution of the modelling equations with a fully implicit Runge-Kutta integration scheme. Finally, optimal profiles of well-known fed-batch fermenter optimization problems are found with this approach. These solutions compare favorably with published analytically-based results and are obtained with modest computational effort Also, imposition of state variable constraints, which cannot be dealt with analytically, is handled here by simply bounding the Lagraage polynomial coefficients. Examples illustrating the general applicability of the method to fed-batch fermentation will also be presented.

We begin the development of a strategy for solution of fed-batch reactor optimal control problems by stating the following general optimization problem. This problem contains both differential and algebraic equations with variables and profiles as decisions.

Min *(x.Utt).ZCt)) (DAOP1)
x,U(t),Z(t)

s.L c(x,U(;t),Z0t)) = 0

g(x,Utt),Z(;t)) * 0

Ut) = F(x,UCt),ZOtU) *t* € [0,1]

Z(0) = Z_o

x^L £ x £ x^u

U^L £ UOt) £ U^u

z^L ≤> *nt)* * z^u

where

4> = an objective function

g,c = design constraint vectors

x = decision variable vector

Tit) = state profile vector

UOt) = control profiles

x^L,x^u = variable bounds

U^L,U^U ss control profile bounds

Z^L,Z^U = state profile bounds

The differential-algebraic optimization problem (DAOP1), as stated, cannot be solved directly by typical nonlinear programming techniques *or* optimal control methods. In general, with an NLP technique one cannot optimize continuous profiles, nor is it possible to impose bounds and/or general constraints involving *Tit)* and UCt). Here even a differential-algebraic equation (DAE) solver can only deal with added algebraic *equality* constrained problems. Coupled with a nonlinear programming technique, repeated, and often expensive, solution of the

DAE model (perhaps including a large number of sensitivity equations) is still required here. Optimal control methods, on the other hand, will optimize continuous control profiles but normally cannot deal with general algebraic constraints (such as c or g).

In the remainder of the paper we present and justify a method for the solution of (DAOP1), especially for biochemical reactor problems. In the next section a discretization is presented which converts the differential equations to a set of algebraic equations. Here orthogonal collocation is applied to finite elements for the purpose of handling discontinuous control profiles. Following this, stability and accuracy properties, which nonnally apply to numerical integration techniques, will be discused as to how they apply to orthogonal collocation. In addition, an equivalence between the Kuhn-Tucker conditions and variational conditions will be derived for a wide class of differential-algebraic optimization problems (DAOP's). The last section deals with the optimization of a fed-batch reactor. In the simplest case it is seen that solutions are found that compare well with an analytically-based solution. More complex cases are then solved and presented which do not have analytic solutions.

2. Discretization of DA0P1 to a Nonlinear Programming Problem

In this section we focus on discretizing the ordinary differential equation (ODE) model of (DAOP1):

1U) = F(x,UU)Mt)X) X€[0,1] (1)

Z(0) = Z_Q .

Here we make use of two polynomial types written in Lagrange form:

$$z_{K+1}(t) = \sum_{i=0}^{K} z_i\, \phi_i(t) \quad \text{where} \quad \phi_i(t) = \prod_{k=0,i} \ldots \qquad (2)$$

$$u_K(t) = \sum_{i=1}^{K} u_i\, \psi_i(t) \quad \text{where} \quad \psi_i(t) = \prod_{k=1,i} \ldots \qquad (3)$$

where the state vector approximation, $z_{..}(t)$, is a (K+1)th order (degree < K+1) polynomial and the control vector approximation, u.,(£), is a Kth order (degree < K) polynomial. Also the notation k = 0,i denotes k starting from zero and k # i. Note that the Lagrange form polynomial has the desirable property that (for z ^ t t) , for example)

**•,<*> - v

Since chemical engineering problems have states and controls that represent quantities like temperature or concentration, using Lagrange polynomials produces coefficients z_i and u_i which are physically meaningful quantities. This becomes useful when providing variable bounds, initializing a profile, or interpreting solution profiles.

Substitution of (2) and (3) into (1) and discretization of the ODE's using orthogonal collocation yields the set of algebraic equations:

$$Rtt_i) = \sum_{j=0}^{K} z_j \, {}^{\wedge}\text{-}tt_i) - F(x,u_i,z_i\, X) = 0 \qquad i=l,...K \qquad (4)$$

with $z_0 = Z_0$.

Using the ODE model (4), (DA0P1) becomes:

Min - 4>(x,t_i,z_i) (NLP1)

x.a,z_i

s.L

g(x,a_i,z_i) £ 0

c(x,a_i,z_i) = 0

$$RO_i) = \sum_{j=0}^{K} z \,.4>U) - F(x,u,z.X) = 0 \qquad i=l....K$$

$z_0 = z_0$

x^L £ x £;x^u

U^L S: a_i £U^U

Z^L * z_i :\wedge Z^u

With (NLP1) we can now solve very general differential-algebraic optimization problems once the points £., i=l....K are chosen. Here the location of these points corresponds to the shifted roots of an orthogonal Legendre polynomial of degree K. The limitation of the formulation in (NLP1) is that all profiles are assumed to be smooth (i.e. analytic functions in £). If this is true, then choosing K sufficiently large will yield accurate solutions. However, in most cases, this assumption does not hold and the formulation must be extended to finite elements.

To extend problem (NLP1) to deal with finite elements we refer to a related paper (Cuthrell and Biegler (1987)) where advantages of finite elements and their appropriate placement is discussed. In that reference, a set of finite element knot placement equations was solved to position the knots in order to minimize the approximation error of the state variable profiles. The problem of control profile discontinuity was also handled in Cuthrell and Biegler (1987) by introducing additional finite elements termed super-elements. The breakpoint (or knot) positions of these elements were included as additional degrees of freedom so that points of discontinuity could be found. To simplify the analysis in this paper, we shall assume that all state profiles can be approximated to suitable accuracy using relatively low order polynomials (e.g. K+l < 5) and do not require finite elements for this purpose. Instead, we focus on using finite elements as super-elements, so that discontinuities of optimal control profiles can be determined through optimal locations of finite element knots.

To preserve the orthogonal properties obtained with global collocation the domain >t£[0,l] is mapped into each finite element through the formula (with a^a, tt_{NE+1}=W:

$$t = a_1 \bullet Hoc._{1+1} - a._1) \qquad i=i....NE \qquad \text{for } t \in [a._1, a._{1+1}].$$

And the locations of the orthogonal Legendre roots (with $t_Q \equiv 0$) are mapped to the points

$$t_{(i-1)(K+1)+j} = \alpha_i + \tau_j(\alpha_{i+1} - \alpha_i) \qquad i=1,\dots,NE \tag{5}$$

It is convenient at this point, in order to save a considerable amount of rewriting, to define the expression (i-l)(K+l)+j by the label [ij]. This label, [ij]≡(i-l)(K+l)+j. is not to be confused with the commonly used double subscripts for matrices (e.g., A_{ij} meaning the element in the ith row and jth column). For an equivalent derivation of finite element collocation which uses matrix notation see Finlayson (1980). Furthermore, the indices "i" and "j" can themselves take on other characters when the context requires it For example, [i—Ik] becomes (i-l)(K+l)+k for some i and k, and [il]=(i-l)(K+l)+l for some i. With this convention (5) becomes:

$$t_{[ij]} = a_i + t(\alpha_{i+1} - a_i) \qquad i=1,\dots,NE \quad j=0,\dots,K. \tag{6}$$

The Lagrange polynomials can now be expressed as:

$$z^i_{K+1}(t) = \sum_{j=0}^{K} z_{[ij]}\,\phi_{[ij]}(t) \qquad \phi_{[ij]}(t) = \phi_j(t) = \prod_{k=0,j}^{K} \frac{(t - t_{[ik]})}{(t_{[ij]} - t_{[ik]})} \tag{7}$$

$$u^i_K(t) = \sum_{j=1}^{K} u_{[ij]}\,\psi_{[ij]}(t) \qquad \psi_{[ij]}(t) = \prod_{k=1,j}^{K} \frac{(t - t_{[ik]})}{(t_{[ij]} - t_{[ik]})} \tag{8}$$

for i=1,...,NE.

The discretized residual or collocation equations can be written down immediately from (4) as:

$$R(t_{[i\ell]}, \Delta\alpha) = \sum_{j=0}^{K} z_{[ij]}\,\dot{\phi}_{[ij]}(t_{[i\ell]}) - F(x,\alpha_i, z_{[i\ell]}, t_{[i\ell]}) = 0 \tag{9}$$

i=1,...,NE
ℓ=1,...,K

with $z_{[10]} = z_0$ •

The calculation of the term $\dot{\phi}_{[ij]}(t_{[i\ell]})$ can be simplified by chain ruling derivatives to obtain:

$$\dot{\phi}(t_{[i\ell]}) = \dot{\phi}(t_\ell) / \Delta\alpha_i \qquad j=0,\dots,K \quad \ell=1,\dots,K.$$

And thus (9) is more simply written as:

$$R(t_{i\ell}) = \sum_{j=0}^{K} z_{ij} \frac{\dot{\phi}_j(t_\ell)}{\Delta\alpha_i} - F(x, u_{i\ell}, z_{i\ell}, t_{i\ell}) = 0 \qquad i=1,\ldots NE \qquad (10)$$

$$\ell=1,\ldots,K$$

with $z_{10} = z_0$.

In (10) the expression $\dot{\phi}_j(t_\ell)$ can be easily calculated offline (see Villadsen & Michelsen (1978)) since it depends only on the Legendre root locations. Now, assuming for the moment the variables x and $u_{i\ell}$ are fixed, (10) is composed of M(NE(K)+1) equations and M(NE(K+1)) state coefficients. To make the system well posed an additional set of M(NE-1) equations is written to make the polynomials z^i_{K+1} (t) continuous at the interior knots α_i, i=2,....NE. This is done by enforcing

$$z^i_{K+1}(\alpha_i) = z^{i-1}_{K+1}(\alpha_i) \qquad i=2,\ldots,NE\ , \qquad (11)$$

or equivalently

$$z_{i0} = \sum_{j=0}^{K} z_{i-1,j}\, \phi_j(t=1) \qquad i=2,\ldots,NE\ . \qquad (12)$$

These equations extrapolate the polynomial z^{i-1}_{K+1}(t) to the endpoint of its element and provide an "initial condition" for the next element and polynomial z^i_{K+1}(t). Each overall approximation to the state profile is therefore a continuous and piecewise polynomial function of order K+1. Stated mathematically, $z_{K+1}(t) \in P_{K+1} \cap C[a,b]$ where P_{K+1} denotes the set of all polynomials of order K+1 and C[a,b] is the set of continuous functions. A number of authors construct *differentiable* and piecewise polynomial approximations, from $z(t) \in P_v \cap C^1[a,b]$, to higher order ODE or PDE systems (Finlayson (1980), Gardini (1985)). However, continuous approximations are sufficient for our case particularly since our examples have discontinuous control profiles and non-differentiable state profiles.

At this point a few additional comments concerning construction of the *control* profile polynomials must be made. Recall that these polynomials use only K coefficients per element and are of lower order than the state polynomials. As a result these profiles are constrained or bounded only at collocation points. Maintaining the entire control profile within the problem constraints is necessary in order to approximate the variational profile better. This can be attempted in many ways. Here we bound the values of each control polynomial at both knots in its element This can be done by writing the equations:

$$U^L \le u_K^i(\alpha_i) \le U^u \qquad i=1,\dots,NE.$$

$$u^L \le u_K^i(\alpha_{i+1}) \le u^u \qquad i=1,\dots,NE.$$

Recall that since control polynomial coefficients exist only at collocation points, enforcement of these bounds can be done by extrapolating the polynomial to the endpoints of the element This is easily done by using:

$$u_K^i(\alpha_i) = \sum_{j=1}^{K} u_{ij}\,\psi_j(t=0) \qquad i=1,\dots,NE.$$

and

$$u_K^i(\alpha_{i+1}) = \sum_{j=1}^{K} u_{ij}\,\psi_j(t=1) \qquad i=1,\dots,NE.$$

Adding these constraints affects only the shape of the final control profile and not the optimal value of the control polynomials at the collocation points. The net effect on these constraints is to keep the endpoint values of the control profile from varying widely outside their ranges $[U^L, U^U]$.

Including the ODE model, discretized on finite elements, the state continuity conditions and bounds on the control profiles at the knots, the NLP formulation becomes:

$$\min_{x,\,u_{i\ell},\,z_{i\ell},\,\alpha_i} \; \phi(x, \alpha_i, u_{i\ell}, z_{i\ell}) \qquad \text{(NLP2)}$$

s.t.

$$g(x, u_{i\ell}, z_{i\ell}) \le 0$$

$$c(x, u_{i\ell}, z_{i\ell}) = 0$$

$$R(t_p) = 0 \qquad i=1,\dots,NE$$

$$\ell=1,\dots,K$$

$$z_{10} = z_0$$

$$z_{i0} = \sum_{j}^{K} z_{(i-1)j}\, \phi_j(t=1) \qquad i=2,\ldots,NE$$

$$u^L \leq u_K^i(\alpha_i) \leq u^U$$

$$u^L \leq u_K^i(\alpha_{i+1}) \leq u^U \qquad i=1,\ldots,NE$$

$$x^L \leq x_{ij} \leq x^U$$

$$u^L \leq u_{ij} \leq u^U$$

$$z^L \leq z_p \leq z^U$$

In this formulation the knot positions, a_i, are formulated as decision variables and found by the optimization procedure as points of control profile discontinuity. Note, however, that the state profiles are required to be continuous because of the continuity constraints.

In the next section we will justify the assumption that the state profiles can be accurately approximated by discussing both stability and accuracy properties of orthogonal collocation. Here we show that solving the differential equations via orthogonal collocation is equivalent to performing a fully implicit Runge-Kutta integration at Gaussian roots.

3. Properties of the NLP Strategy

Before applying the NLP strategy to the *fed-batch fermenter example it is useful to consider a few properties of the method proposed above. The following subsections therefore address the following questions:

- **Given a control profile, can the differential equation model be solved accurately and stably using orthogonal collocation on finite elements?**

- **How does the DAOP solution found with nonlinear programming compare to a solution that solves the variational conditions of the optimal control problem? How do the optimal control profiles compare?**

3.1. Stability and Accuracy of Orthogonal Collocation

In this section additional theoretical properties of orthogonal collocation will be considered. First, we will see that, under certain conditions, the method of orthogonal collocation on finite elements is equivalent to performing a *fully implicit* Runge-Kutta (RK) integration of the ODE's at Gaussian points. The conditions required for this are, simply stated, that collocation be done at Gaussian roots, and that the elements in the Butcher's block array be specified in a certain way for the Runge-Kutta method. In establishing the equivalence between our method and a RK method, a number of important theoretical properties, normally associated only with numerical integration schemes, can immediately be applied. These properties include stability, symmetry, and the order of the method, the theoretical analysis cited here is based on the work of Burrage and Butcher (1979) and Ascher and Bader (1986).

First, for a numerical integration method to be considered stable, it must result in integrations of an ODE which differ by no more than the difference in initial conditions (see Burrage and Butcher (1979)). An integration method is said to be symmetric if it is invariant (or gives the same solution) under a change in the direction of integration. The concept of symmetry is normally unimportant for integration of initial-value problems becomes quite important when solving boundary-value problems.

To define the stability of a particular ODE method, a test equation is usually chosen for analysis. For example, the standard test function

$$\dot{y}(t) = X \; y(t) \tag{13}$$

where X is a complex number with a nonpositive real part,

is used to define A-stability. It is well-known that implicit linear multi-step methods of orders £ 2 (Dahlquist (1963)), and all fully implicit RK methods of order 2K (Ehle (1968)), where K is the number of function evaluations required per step, are A-stable. Simply stated, the term "A-stable" implies that, for all Re(X)£0, these integration schemes, stably integrate (13).

With X in (13) replaced by X(t) the test function is said to be nonautonomous and the corresponding stability property is termed AN-stability. For this case X(t) is considered a function having values in the nonpositive complex plane.

Generalizing further, numerical integration schemes are said to possess B-stability if they stably integrate the nonlinear autonomous system:

$$\dot{y}(t) \bullet F(y(t))$$

$$\text{if } (F(y)-F(z))^T(y-z) \; \pounds \; 0 \quad V \; y.z,$$

and, BN-stability if they stably integrate the nonlinear nonautonomous system:

$$\dot{y}(t) = F(t,y(t)) \tag{14}$$

$$\text{if } (F(t,y)-F(t,z))^T(y-z) \; * \; 0 \quad V \; y,z,t$$

The restrictions on the right hand sides here merely imply that F is a monotonic nonincreasing function, and are necessary for y(t) to approach some steady state value.

Burrage and Butcher (1979) prove that implicit RK methods which possess a property known as *algebraic* stability also possess the properties of A, AN, B, and BN-stability. Or, simply stated, * implicit RK methods which are algebraically stable, will stably integrate autonomous and nonautonomous, linear and nonlinear systems which satisfy the above restrictions. Algebraic stability is thus a stronger condition than the other types of stability and does not depend on any particular test function. In fact it derives from a matrix property that involves only the values in the Butcher's block array (see Burrage and Butcher (1979)). Although this does not guarantee stability for all systems of ODE's, it does indicate that the method is stable for a wider class of problems (i.e., (14)) than previously shown.

Later, Ascher and Bader (1986), constructed several fully implicit, symmetric, algebraically stable RK schemes and showed that the only equivalent methods to these are finite element collocation methods which use Gaussian roots. An example that illustrates this equivalence can be found in Cuthrell (1986). Finally, Ascher and Bader (1986) also showed that A-stable integration schemes may not be suited for solving BVP's and. that, instead, the stronger condition of algebraic stability is required. This property is important here because optimal control problems usually require ODE solutions in both the forward direction (for the model equations) and the backward direction (for the adjoint equations). In the next subsection we will see that the method proposed above solves both equations sets simultaneously by solving (NLP2).

3.2. Analysis of the Optimality Conditions

In the previous subsection we were concerned about the accuracy of the state variable profile with control profiles specified. Here we consider, instead, how accurately control profiles are determined with nonlinear programming. In (NLP2) optimization is done over the variables x, and the profiles $z_{K+J}(t)$ and $u_K(t)$, while the original problem, (DAOP1) involves the variables x and the profiles Z(t) and U(t). We now consider the conditions for the continuous profile U(t). Since the analysis follows the same lines for the parameters, x, and the optimality conditions are similar, we will not be concerned with these parameters in this subsection.

In order to present a discussion of the accuracy of the optimal solution, we begin with the following general DAOP (DAOP2) and write the nonlinear program resulting from application of orthogonal collocation, we then write the Kuhn-Tucker conditions for this problem and show that these can be written as discrete analogs of the variational conditions of (DAOP2). For this analysis, the differential-algebraic optimization problem will be given as:

$$\min_{U(t),Z(t)} \Psi(U(b),Z(b)) + \int_a^b \Phi(Z(t),U(t))\,dt \qquad \text{(DAOP2)}$$

$$\text{s.t.} \quad c(U(t),Z(t)) = 0$$

$$g(U(t),Z(t)) \le 0$$

$$\dot{Z}(t) = F(U(t),Z(t))$$

$$g_f(Z(b)) \le 0$$

$$c_f(Z(b)) = 0$$

$$Z(a) = Z_0$$

Now, by approximating the state variable profiles as continuous, piecewise polynomial functions of order K+1 and control variable profiles as piecewise continuous, polynomial functions of order K, we can apply orthogonal collocation on finite elements and obtain the following nonlinear program.

$$\min_{u_{ijl},\, z_{ijl}} \Psi(z_f) + \sum_{i=1}^{NE} \sum_{j=1}^{K} M_{ij}\, \Phi(z_{ij}, u_{ij}) \qquad \text{(NLP3)}$$

$$\text{s.t.} \quad \dot{z}_{K+1}(t_{ij}) - F(z_{ij}, u_{ij}, t_{ij}) = 0$$

$$g(z_{ij}, u_{ij}) \le 0$$
$$c(z_{ij}, u_{ij}) = 0$$
$$g_f(z_f) \le 0$$
$$c_f(z_f) = 0$$
$$z_{10} = z_0$$
$$z_{i0} - z^{i-1}_{K+1}(\alpha_i) = 0 \qquad i = 2, \ldots, NE$$

To simplify the analysis we omit the extrapolation constraints on the control profiles, although as seen in the previous section their inclusion does not change the results. We also include the extrapolation of $z_{K+1}(t_f)$ as a new variable, z_f. Note that the integral in (DA0P2) is now approximated by a Gaussian quadrature formula summed over all of the elements. Now, on a simple optimal control problem without algebraic constraints, Reddien (1978) showed the equivalence between variational conditions for the optimal control problem and the optimality conditions for the resulting nonlinear program. In his analysis, B-spline basis functions were used on finite elements. Here we adapt and extend his approach by considering Lagrange basis polynomials and the more complicated problem given by (DAOP2) and (NLP3). First, we form the Lagrange function for (NLP3):

$$L(\cdot) = \sum_{i=1}^{NE} \sum_{j=1}^{K} \omega_{ij} \Big[\Phi(z_{ij}, u_{ij})$$
$$+ \lambda_{ij}^T \big(F(z_{ij}, u_{ij}) - \dot{z}_{K+1}(t_{ij})\big) + \mu_{ij}^T\, g(z_{ij}, u_{ij})$$
$$+ \nu_{ij}^T\, c(z_{ij}, u_{ij}) \Big] + \Big\{ \Psi(z_f) + \mu_f^T g_f(z_f) + \nu_f^T c_f(z_f) \Big\}$$
$$+ \gamma_1^T (z_{10} - z_0) + \sum_{i=2}^{NE} \gamma_i^T \Big(z_{i0} - z^{i-1}_{K+1}(\alpha_i) \Big)$$
$$+ \gamma_{NE+1}^T \Big(z_f - z^{NE}_{K+1}(\alpha_{NE+1}) \Big) \qquad (15)$$

Note that we have defined the Kuhn-Tucker multipliers (X_f *fl* and *V)* differently than usual, but equivalently, by including the positive quadrature weights, vo_{ij}. In addition, we assume that the adjoint variable profile can be approximated as $X_{K+1}(t)$, a continuous, piecewise polynomial function of the same order as $z_{K+1}(t)$ with $X_{K+1}(t_{irj})$ = X_{ij} . Now since z^ is a piecewise polynomial of order K, we can make the following equivalence by noting that K-point Gaussian quadrature is exact for polynomial integrands of degree 2K - 1 or less (see, e.g. Carnahan, Luther and Wilkes (1969)). Thus for the quadrature terms in (15) we can write the following relation:

$$-\sum_{i=1}^{NE}\int_{\alpha_i}^{\alpha_{i+1}} \lambda_{K+1}^T \dot{z}_{K+1}\, dt = -\sum_{i=1}^{NE}\sum_{j=1}^{K} \omega_{ij}\, \lambda_{ij}^T\, \dot{z}_{K+1}(t_{ij}) .$$

Integrating by parts yields:

$$-\int_{\alpha_i}^{\alpha_{i+1}} \lambda_{K+1}^T \dot{z}_{K+1}\, dt = -\lambda_{K+1}^T(\alpha_{i+1}) z_{K+1}(\alpha_{i+1}) + \lambda_{K+1}^T(\alpha_i) z_{K+1}(\alpha_i) + \int_{\alpha_i}^{\alpha_{i+1}} \dot{\lambda}_{K+1}^T z_{K+1}\, dt.$$

Writing the quadrature formula for the integral on the right hand side and substitution into (IS) yields the following Lagrange function:

$$L = \sum_{i=1}^{NE} \sum_{j=1}^{K} \omega_{ij} \Big[\Phi(t_{ij}) + \lambda_{ij}^T F(t_{ij}) +$$

$$z_{ij}^T \dot{\lambda}_{K+1}(t_{ij}) + \mu_{ij}^T\, g(z_{ij}, u_{ij})$$

$$+ \nu_{ij}^T\, c(z_{ij}, u_{ij}) \Big] + \Big\{ \Psi(z_f) + \mu_f^T g_f(z_f) + \nu_f^T c_f(z_f) \Big\}$$

$$- V] \ \bullet \sum_{i=1}^{NE} (\gamma_i + \lambda_{K+1}(\alpha_i))^T z_{[i0]}$$

$$- \sum_{i=2}^{NE} (\gamma_i + \lambda_{K+1}(\alpha_i))^T z^M_{K+1}(\alpha_i) + \gamma^T_{NE+1} z_f$$

To simplify this function we first consider the optimality conditions for the continuity variables $z_{[i0]}$. By noting that, from (11) and (12),

$$z^{i-1}_{K+1}(\alpha_i) = \sum_{j=0}^{K} z_{[i-1,j]} \phi_j(t=1) \qquad i=2,\dots,NE . \qquad (12)$$

we have the following set of equations:

$$\# \quad -\Leftarrow r, ^+ \ \backslash_{+1} te,» \ - \ \wedge_{1+1} + x_{K+i}(a_{i+i}» \ \wedge_0 a) = 0 \qquad i = i....NE$$

HOI

Since this system of equations is overdetermined and has an infinite number of solutions, we force a solution by making the assignment that $\gamma_i + \lambda_{K+1}(\alpha_i) = 0$. With this simplification the Lagrange function becomes:

$$L = \sum_{i=1}^{NE} \sum_{j=1}^{K} \omega_{[ij]} [\ \Phi(t_{[ij]}) \bullet \lambda^T_{[ij]} F(t_{[ij]}) +$$

$$z^T_{[ij]} \lambda_{K+1}(t_{[ij]}) + \mu^T_{[ij]} g(z_{[ij]}, u_{[ij]})$$

$$+ \nu^T_{[ij]} c(z_{[ij]}, u_{[ij]}) \ ] + \{ \ \Psi(z_f) + \mu^T_f g_f(z_f) + \nu^T_f c_f(z_f) \ \}$$

$$+ \lambda_{K+1}(\alpha_1)^T z_0 - \lambda_{K+1}(\alpha_{NE+1})^T z_f$$

By noting that $\omega_{[ij]}$ are nonzero, the remaining Kuhn-Tucker conditions can now be derived immediately:

(a) $$\frac{\partial\Phi(t_{[ij]})}{\partial z_{[ij]}} + \left(\frac{\partial F(t_{[ij]})}{\partial z_{[ij]}}\right)\lambda_{[ij]} + \dot{\lambda}_{K+1}(t_{[ij]})$$

$$+ \left(\frac{\partial g}{\partial z_{[ij]}}\right)\mu_{[ij]} + \left(\frac{\partial c}{\partial z_{[ij]}}\right)\gamma_{[ij]} = 0$$

$$\lambda_{K+1}^{i-1}(\alpha_i) = \lambda_{K+1}^{i}(\alpha_i) \qquad i = 2\ldots NE$$

(b) $$\frac{\partial\Phi(t_{[ij]})}{\partial u_{[ij]}} + \left(\frac{\partial F(t_{[ij]})}{\partial u_{[ij]}}\right)\lambda_{[ij]} + \left(\frac{\partial g}{\partial u_{[ij]}}\right)\mu_{[ij]} + \left(\frac{\partial c}{\partial u_{[ij]}}\right)\gamma_{[ij]} = 0$$

(c) $$\frac{\partial\Psi}{\partial z_f} + \frac{\partial g_f}{\partial z_f}\mu_f + \frac{\partial c_f}{\partial z_f}\gamma_f - \lambda_{K+1}(\alpha_{NE+1}) = 0$$

(d) $$\lambda_{K+1}(\alpha_1) = 0 \qquad \text{(if } Z_0 \text{ not specified)}$$

(e) $$\dot{z}_{K+1}(t_{[ij]}) - F(z_{[ij]}, u_{[ij]}, t_{[ij]}) = 0 \qquad z_{[10]} = Z_0$$

$$z^{i-1}_{K+1}(\alpha_i) = z^{i}_{K+1}(\alpha_i) \qquad i = 2,\ldots,NE$$

$$z_f = z_{K+1}^{NE}(\alpha_{NE+1})$$

(f) $$g(z_{[ij]}, u_{[ij]}) \le 0$$

$$c(z_{[ij]}, u_{[ij]}) = 0$$

(g) $$C_f(Z_f) = 0$$

$g_f(z_f) \leq 0$

(h) $\mu_f, \mu_{ij} \geq 0$

(i) $\mu_{ij}^T(g(z_{ij}, u_{ij})) = 0$

$\mu_f^T g_f(z_f) = 0$

We now see that (a) is a discrete analog of the adjoint equations which can easily be derived from standard optimal control theory. They can also be obtained diTectly by applying orthogonal collocation on finite elements to the adjoint equations. Similarly, (b) is a discrete analog of the variational conditions on the control profiles while (c) and (d) represent relationships on the final and initial conditions, respectively, for the adjoint variables. Finally, (e) through (g) are simply feasibility conditions for the ODE's and the problem constraints and the remaining expressions relate to the optimality conditions for the inequality constraints.

Note that, except for the conditions on purely state variable constraints (e.g., g(Z(t))£0), all of the above equations are discrete analogs of the conditions found in Bryson and Ho (1975) for variational problems. To deal with state variable constraints, variational conditions have traditionally been defined using higher time derivatives of these constraints in order to express them as functions of the control. This approach has often been used in order to facilitate analytic solution of the optimal control problem. On the other hand, Jacobson et al (1971) derived alternate conditions for state variable constraints. Here, consider the problem (SCP1) with purely state variable constraints:

Min$_{U(t)}$ v £ *(Z(t),U(t)) dt (SCP1)

S.L tit) = F(Z(t),U(t)) 2Xa) = Z_0

g(Z(t)) * 0.

The conditions derived by Jacobsen et al. (1971) for this problem are:

(a) $\frac{\partial \Phi}{\partial U} + \frac{\partial F}{\partial U} \Lambda = 0$

(b) $\frac{\partial \Phi}{\partial Z} + \left(\frac{\partial F}{\partial Z}\right)$ A + (- | |) M + A(t) = 0 A(b) = 0

(c) g(Z(t)) <£ 0

(d) M(t) g(Z(t)) = 0, M(t) ^ 0

(e) *lit)* = F(Z(t),U(t)) Z(a) = Z_Q

where M(t) and A(t) are adjoint functions for the constraint g(Z(t)) £ 0, and the ODE model respectively.

Note that these variational conditions can now be related to the Kuhn-Tucker conditions presented above. Moreover, Kreindler (1982) showed that the above equations are, in fact, stronger necessary conditions for optimally of (SCP1) than those presented in Bryson and Ho (1975). Thus, we have shown the similarity between the solution solved with a nonlinear-programming formulation and the corresponding variational conditions of the optimal control problem, (DAOP2).

Finally, Reddien (1978) showed on a simpler optimal control problem *without* algebraic constraints that

- The solution of the discrete approximation of the necessary conditions converges to the continuous solution as the level of approximation increases, i.e., as K-»oo. Also, the rate of convergence is of order 2K *if* the functions • and F are K + 1 times differentiable in Z and U, and K times differentiate in t

- Solving a discrete approximation of this problem using an NLP, is equivalent to solving a discrete version of the necessary conditions.

- And, thus, NLP solutions of a discretized problem approach the actual solution as K->oo.

From these conditions one can argue qualitatively that as the level of approximation increases, solutions of (NLP3) will also approach the solution for (DAOP2). However, this will be true only if the state and control profiles are at least continuous within each element. i.e, all profile discontinuities are located at the element breakpoints. Since optimal locations of these discontinuities are generally unknown *a priori,* we must therefore determine their locations by including the element breakpoints as degrees of freedom in (NLP3).

4. Solution of A Fed-Batch Fennenter Problem

Finally, we consider the problem of finding an optimal time-varying control strategy for a fed-batch reactor problem. This is a difficult optimal control problem involving the biosynthesis of penicillin. The problem solved here has been treated in a series of papers by Modak et al. (1986), Lim et al. (1986) and Tayeb and Lim (1986). We first state the problem and outline some of its features. Next the technique proposed above is applied for cases with known solutions as well as more complex ones.

Consider Figure 4-1 which presents a schematic diagram of a fed-batch reactor. Here, the reactor contains biomass (X). substrate (S) and product (P) at certain concentrations (grams/liter), and has volume (V) (liters). The control profile for this system is the feed rate (U) (grams/hour) of substrate.

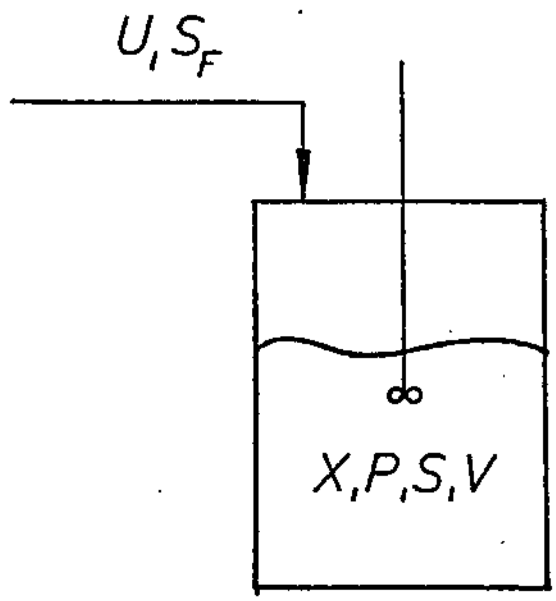


Figure 4-1: Diagram of A Fed-Batch fennenter

The differential-algebraic optimization problem is stated below in problem (BFP). The model and associated kinetic parameters were taken from Lim et al. (1986) (Figure 2 p. 1415). (BFP) is a difficult DAOP because of the state prof Ue and control profile inequality constraints (i.e. profile bounds) and the linear dependence on the control. For problems linear in the control the optimal control profile is either bang-bang or contains singular control arcs. For singular arcs, the control profile does not directly influence the optimality conditions of the Hamiltonian and thus determining U(t) requires additional conditions that can be difficult to handle. As we shall see, the solution to (BFP) does in fact have singular control arcs.

The fed-batch fermenter problem can be stated as:

$$\min_{U(t),T} \wedge \bullet = -P(T)V(T) \qquad \text{(BFP)}$$

s.t.

X(t) = *IKX.S)* X - (- ^ y) U (subscript F) X(0) = 1.5 *g/l*

Pit) = *P(S)* X - K_{degr} P - (- ^ V) U (subscript F) P(0) = 0.0 g/ℓ

S(t) = - M(x,s) (* / $Y_{x/s}$) - *p(s)* (X / $Y_{p/s}$)

- (m S / K_m W) X x + (1S/ V V

S(0) = 0.0 g/ℓ

V̇(t) = U/S_p V(0) = 7 ℓ

M(X,S) = \L (,— jj— -) x

ρ(S) = max(K AS+ b ') (subscripts p, in)

0 £ X(t) £ 40 %/L   \max\ K + S(l+S/K) /

0 £ S(t) £ 100 %/L

0 £ V(t) £ 10 £

0 ≤ U(t) £ 50 g S/hr

72 £ T £ 200 hr

where A(X,S) = growth rate of biomass (hr)*[1]

P(S) = production rate of penicillin (g P/g X-hr)

μ_{max} = 0.11 hr^{-1}

ρ_{max} = 0.0055 g P/g X-hr

K_x = 0.006 g S/g X

K_P = 0.0001 g S/l

K_{in} = 0.1 g S/l

K_{degr} = 0.01 hr^{-1}

K_m = 0.0001 g S/l

m_s = 0.029 g S/g X-hr

$Y_{X,S}$ = 0.47 g X/g S

$Y_{P/S}$ = 1.2 g X/g S

S_F = 500 g S/l

It is important to point out that two differences exist between the above formulation and that cited in Lim et al. (1986). Both differences are due to inconsistencies in the Lim et al. (1986) paper (Modak (1987)). First the maintenance term in the substrate equation was changed from 0.029X (cf. (23) p. 1414 Lim et al. (1986)) to 0.029S/(K_{in}+S). Second the value of 0.004 (cf. (22) ibid.) was corrected to 0.0055 for ρ_{mix}. It should be noted that the above corrections must be made (Modak (1987)) in order to reproduce the results given in Figure 2 of Lim et al. (1986). For more details about this fermenter problem the interested reader is also referred to Bajpai and Reuss (1980, 1981).

4.1. An **Analytically-Based** Solution of Problem **(BFP)**

The fed-batch fermenter control problem stated in (BFP) has been solved in Modak et al. (1986) and Lim et al. (1986). As explained below, this solution derives from the variational conditions, but also requires repeated numerical solution of the fermenter model. We reproduced the solution profiles of this problem and present them in Figures 4*2 to 4-4 The values for the points of control profile discontinuity, final time and the optimal value of the objective function are 11.21, 28.79 and 124.9 and -87.05, respectively. These values can be found in Lim et al. (1986), except the value of the objective function results from integrating their model with the changes indicated above. Lim et al. (1986) quote an objective function value of -86.99 (i.e. a penicillin yield of 86.99 g).

4.2. Four Numerical Solutions of Problem (BFP)

To demonstrate the method proposed above, four numerical solutions of (BFP) are presented. Each was solved using the NLP structure given in (NLP2) and the SQP algorithm presented in Biegler and Cuthrell (1985).

Initialization of the NLP starting point was done as follows. First a continuous control profile U(t) was used along with model initial conditions and initial batch time, T, to integrate the model using the ODE solver LSODE (Hindmarsh (1980). Next the number of collocation points K and the number and locations of the finite element knots were chosen, the collocation point locations were next calculated based on (6). State profile polynomial coefficients at the collocation points were then initialized from these continuous profiles. Polynomial coefficients at the knots were computed by solving the linear system of continuity equations, (12). Since these equations are linear, solving them beforehand keeps the optimization algorithm in the feasible subspace of these constraints.

In Case I (BFP), as stated above, was solved. Here the optimal control profile given in Figure 2 of Lim et al. (1986) (or Fig. 4-2 here) was used as the basis for the starting point Also three finite elements were used with the initial breakpoint positions at 11.21 and 28.79 and a final (or batch) time of 124.9. In Case II an initial constant control profile of U(t)=25.0 and a general initial knot distribution and batch time (see Table 4-1) were used. In Case III, the problem was modified by changing the substrate upper bound to S(t) £ 25.0. Finally, in Case IV a more complex objective function was used that includes the penicillin yield and batch time as well as the cost of the feed solution, and.reflects the net profit of the process per batch:

$$\Phi = -2.5X10^{-2} P(T)V(T) + 168T + 8.5XHT \int_0^T V_f U(t)\, dt$$

Table 4-1 lists the number of elements (NE), collocation points (K), initial knot

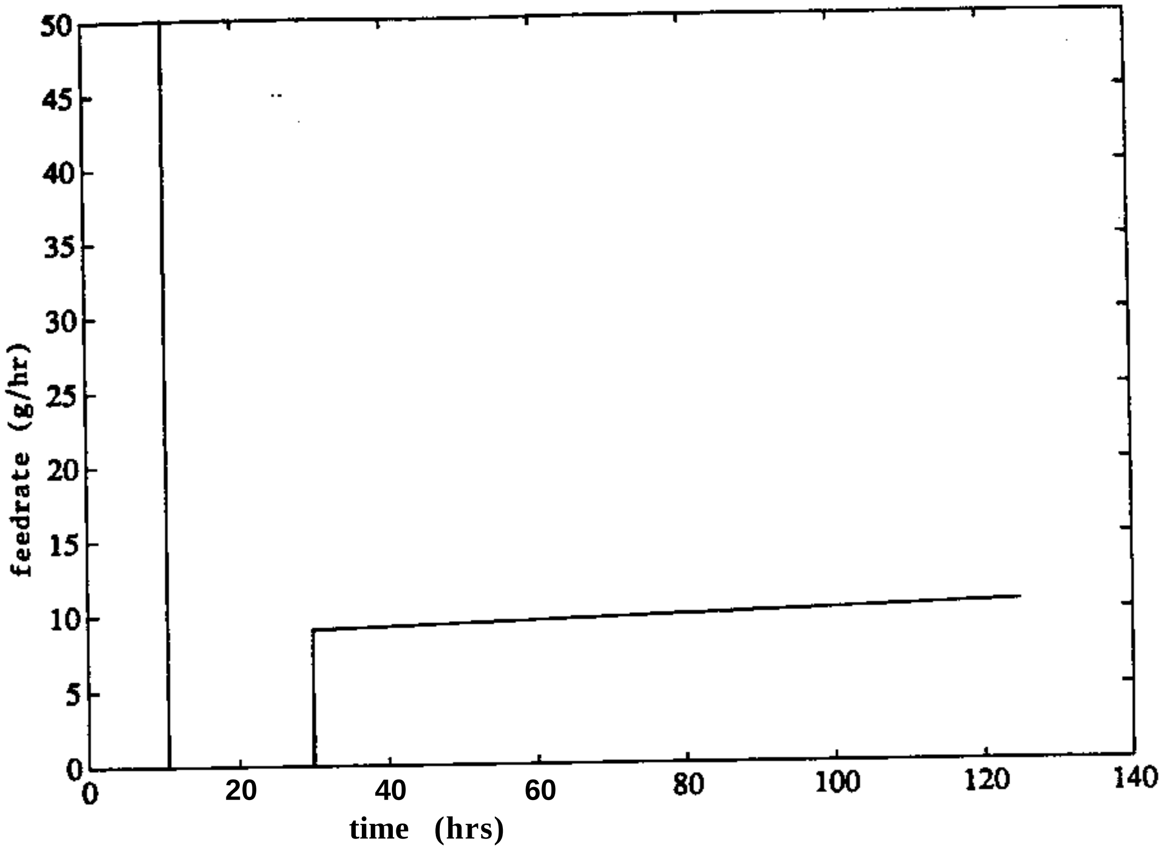


Figure 4-2: Analytically-Based Control Profile

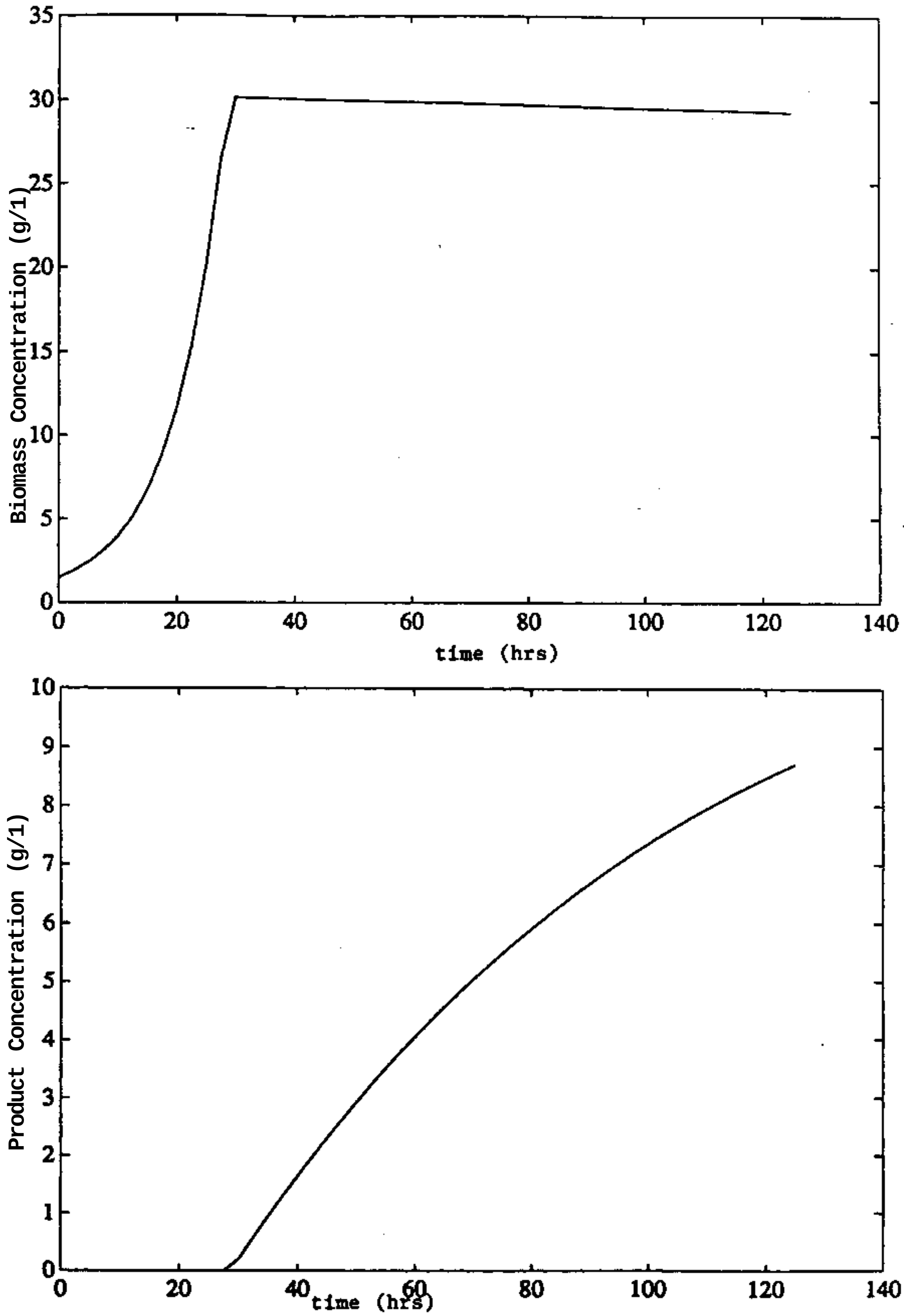


Figure 4-3: Analytically-Based Biomass and Product Profiles

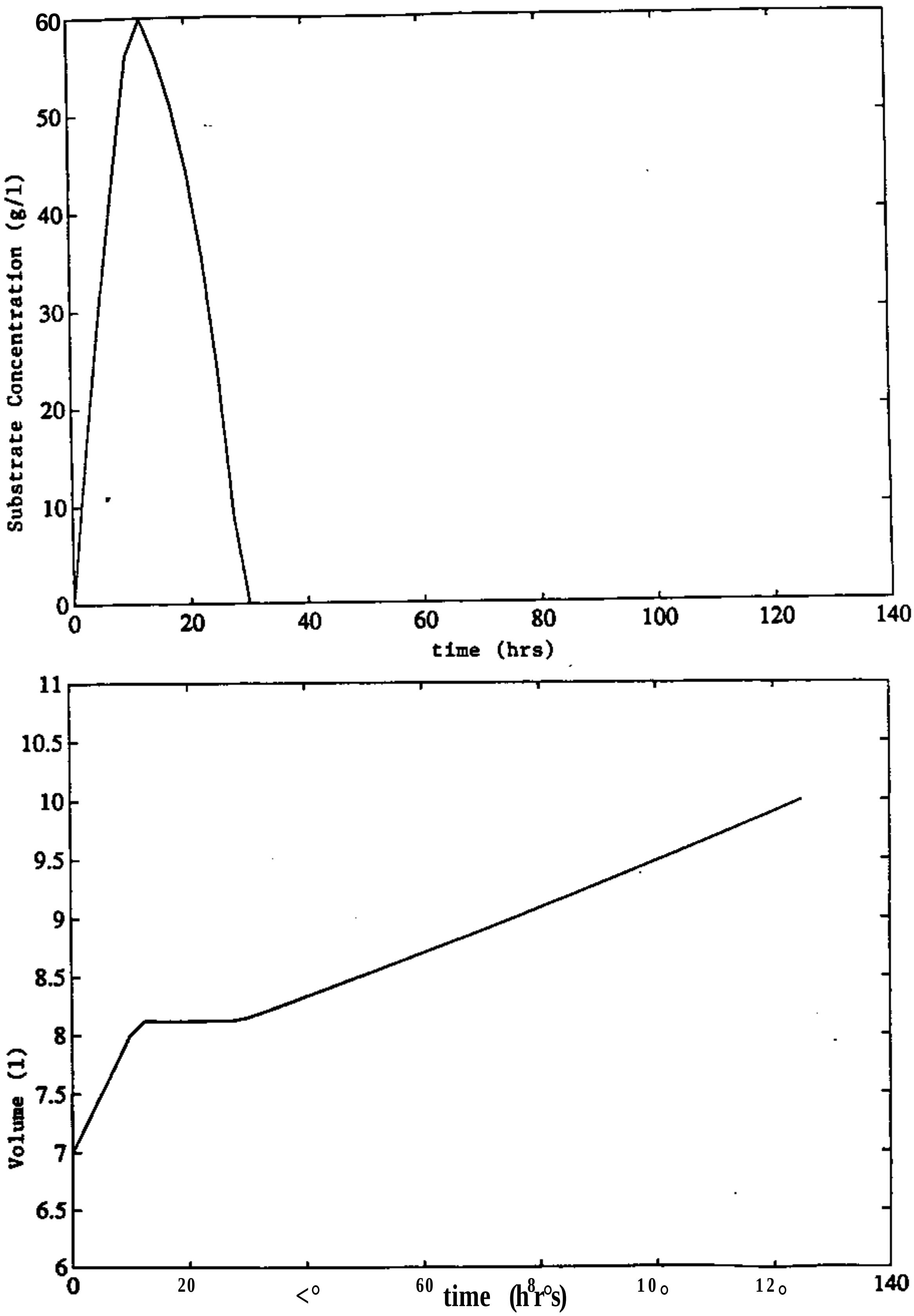


Figure 4-4: Analytically-Based Substrate and Volume Profiles

distribution used a^0 for the above cases as well as the initial batch time T° and the initial objective function value 4>°.

Case	NE/K	a^0	T^0	Φ^0
I	3/4	11.21,28.79	124.9	-82.90
II	3/3	20,50	80	-73.07
III	3/4	20.50	80	-73.98
IV	3/3	20,40	80	134.39

Table 4-1: Initial Parameters and Objective Function Values

The results for the three cases are presented in Table 4-2 and Figures 4-5 through 4-16. In Table 4-2, # iter refers to the number of QPs required to solve the NLP, KT error is the final Kuhn-Tucker error and T* and •* are the optimal batch time and objective function. The optimal knot distribution is denoted by a*. The convention A(-b)=AX10"^{b} has also been adopted. In the figures, all control profiles are drawn with solid lines and depict polynomials constructed with the Lagrange formula (8). Note that the controls have no initial conditions and are not required to be continuous at the knots. The state profiles are drawn with broken lines, if they are approximations and solid lines if they were obtained from an *a posteriori* integration of the model.

By comparing the optimal profiles (Figs. 4-5,6,7) and parameters (Table 4-2) obtained in Case I it can be seen that solutions found using the proposed math programming approach are close to the analytically-based solution. Here slight differences can be seen in the math programming results. For example, the control profile is not at its bounds during the first two feeding periods, the substrate profile attains a slightly lower peak value, and the, biomass profile has a slightly altered slope after ~30 hrs. However, in both cases the volume is at its upper bound at final time and the product profiles are almost identical. Also, the optimal objective function value, final time value and knots for the NLP solution differ only by a small amount: -87.83, 128.29 and 11.46 and 29.29 vs. -87.05, 124.9 and 11.21 and 28.79 from

Case	# iter	KT error	a^*	T_f^*	Φ^*
I	25+	1.26(-2)	11.46,29.29	128.29	-87.83
II	31+	2.05C-2)	29.01,29.38	126.34	-87.79
III	32f	5.88(-2)	21.27,32.60	132.14	-87.69
IV	8	7.02C-8)	18.54,28.94	72.0	120.95

t terminated due to two line search failures

Table 4-2: Results for (BFP)

the Lim solution. This represents an *improvement* in the objective function over the analytically-based solution. This improvement is worth noting because the state variable profiles, as seen in Figs 4-6 and 4-7, satisfy the model *exactly.*

To explain this improved solution, even though the NLP control profile does not exhibit bang-bang features, we need to consider the approach of Lim et al. Using variational conditions, they were only able to predict the correct *shape* of the control profile (bang-bang-singular arc). However, a trial and error search was still required to determine the optimal switching times, and this may not lead to the exact optimum. On the other hand, note that the NLP solutions for Cases I—III all have singular arcs very similar to Lim's solution, even though the profiles are quite different *before* the singular arc is encountered. Because the objective functions are almost identical, however, this seems to suggest that the optimal control profile before the singular arc may be nonunique.

In Cases II and III a starting profile of U(t)=25.0 is used with initial knots at 20 and 50. In Case II some similarities can been seen with the results of Case I. Most importantly the optimal objective function value of -87.79 and the final time of 126.34 are close to the analytically-based values as well as those in Case I. Both the biomass and product profiles compare favorably to the results in Case I, but differences can be seen in the control and substrate profiles. The control profile appears to be just two singular arc portions with the knots being very close together at 29.01 and 29.38. The substrate profile is different in that only a maximum of about 30 *g/i* is achieved. Very small differences between the profiles obtained with Lagrange polynomials and those found via numerical ODE solution with LSODE are also noticeable and are indicated by the dotted and solid lines.

Given the results of both Cases I and II it is clear that the optimum for (BFP) is relatively insensitive to the final control profile. Observe the differences between the analytically-based profiles and knot positions, and those in Cases I and II. Yet for these cases the optimal value of the objective function differs by only 0.04 g. This relative insensitivity of the objective function results in nonunique solutions being obtained and also in some numerical convergence difficulties. It is clear that optimal knot positions, final times and the final control profile vary measurably from the analytically-based results. Tight convergence of the Kuhn-Tucker tolerance was also difficult to attain; tolerances smaller than 0.01 could not be obtained before line search failures occurred. Note that in Case IV much tighter convergence is attained with a different objective function. However, both Cases I and II have objective function values that are noticeable improvements over the Lim solution and have reasonably accurate state variable profiles as well.

The problem posed in Case III is different than those of Case I and II. Here we have additionally imposed the state variable constraint S(t)£25.0 by simply changing the upper bound S^u to 2S.0. The results obtained are presented in Table 4-2 and Figs. 4-11 to 4-13. Good accuracy of the state profile approximation again results with very minor differences obtained between the polynomials and integrated profiles. Case HI illustrates the flexibility of the NLP method over an analytically-based approach. In the NLP approach additional general constraints (not necessarily limited to the substrate profile constraint we used) can be imposed with little difficulty. This is not true of other methods such as that of Lim et al. (1986). Here we easily enforced S(t)£25.0 and obtained both good results in terms of the optimal control profile and the state profile approximation accuracy.

Finally, in Case IV use of a different objective function results in an optimal batch time of 72 hrs (its lower bound) and optimal knots at 18.54 and 32.60. Observe again (Figs. 4-14 to 4-16) that excellent results are obtained in terms of the accuracy of the state profiles, and a final Kuhn-Tucker error of 7.02X10"^{8} was obtained. Note that in this case the final volume is not at its upper bound.

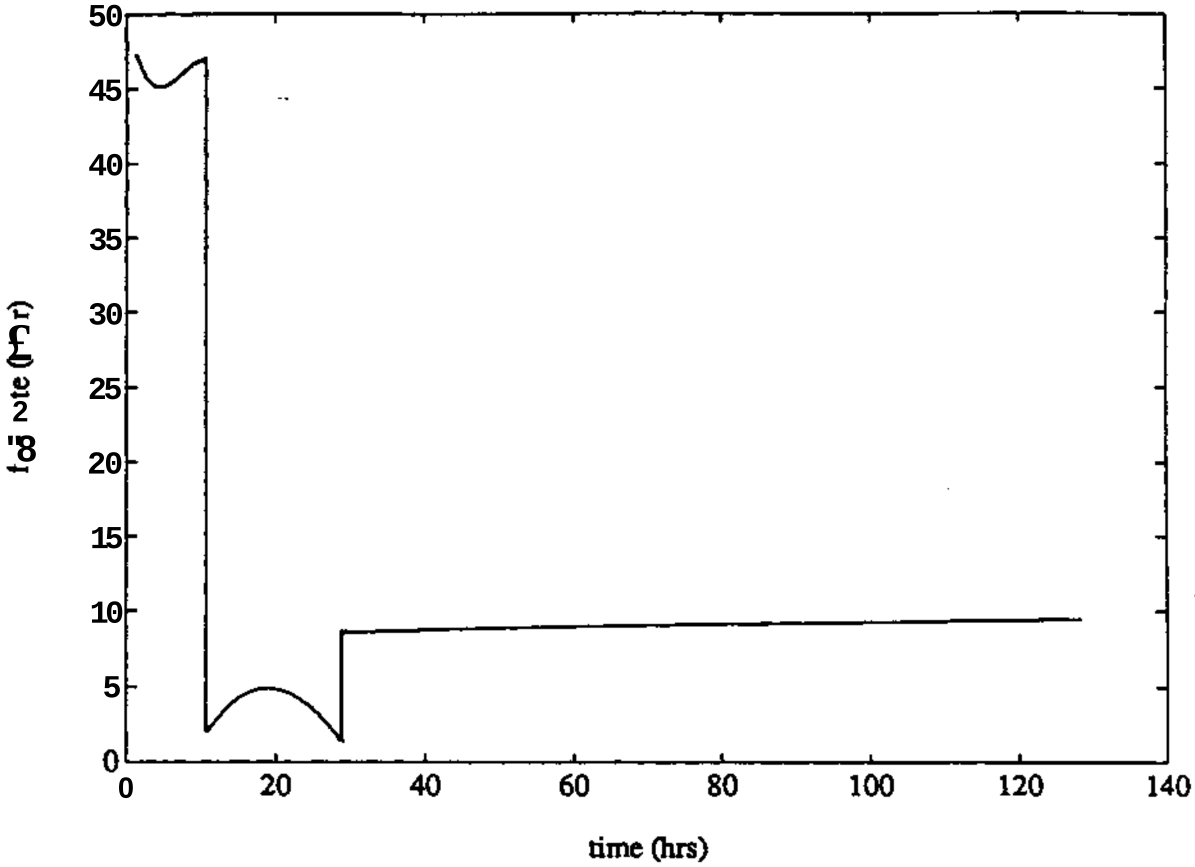


Figure 4-5: Control Profile for Case I

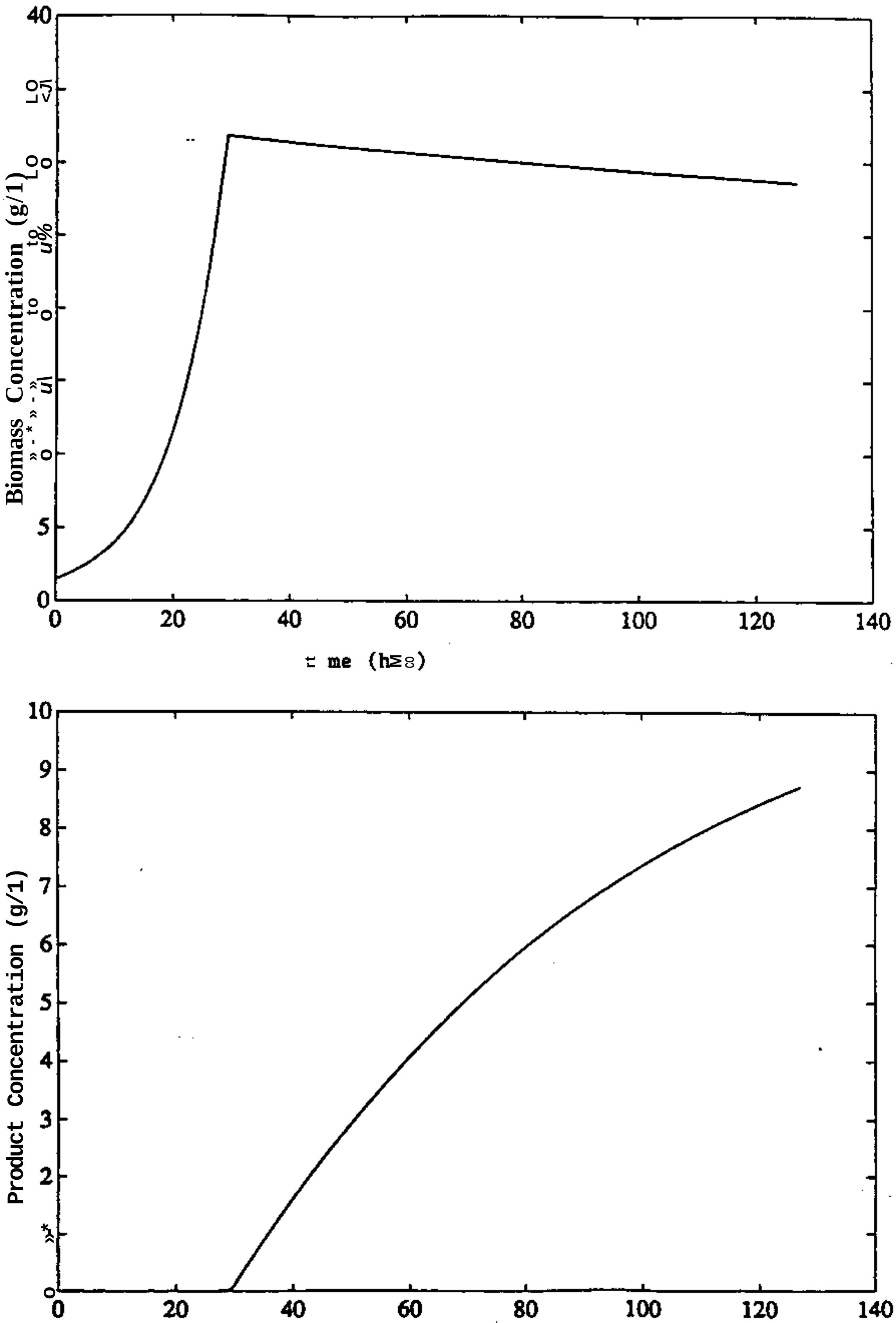


Figure 4-6: Biomass and Product Profiles for Case I

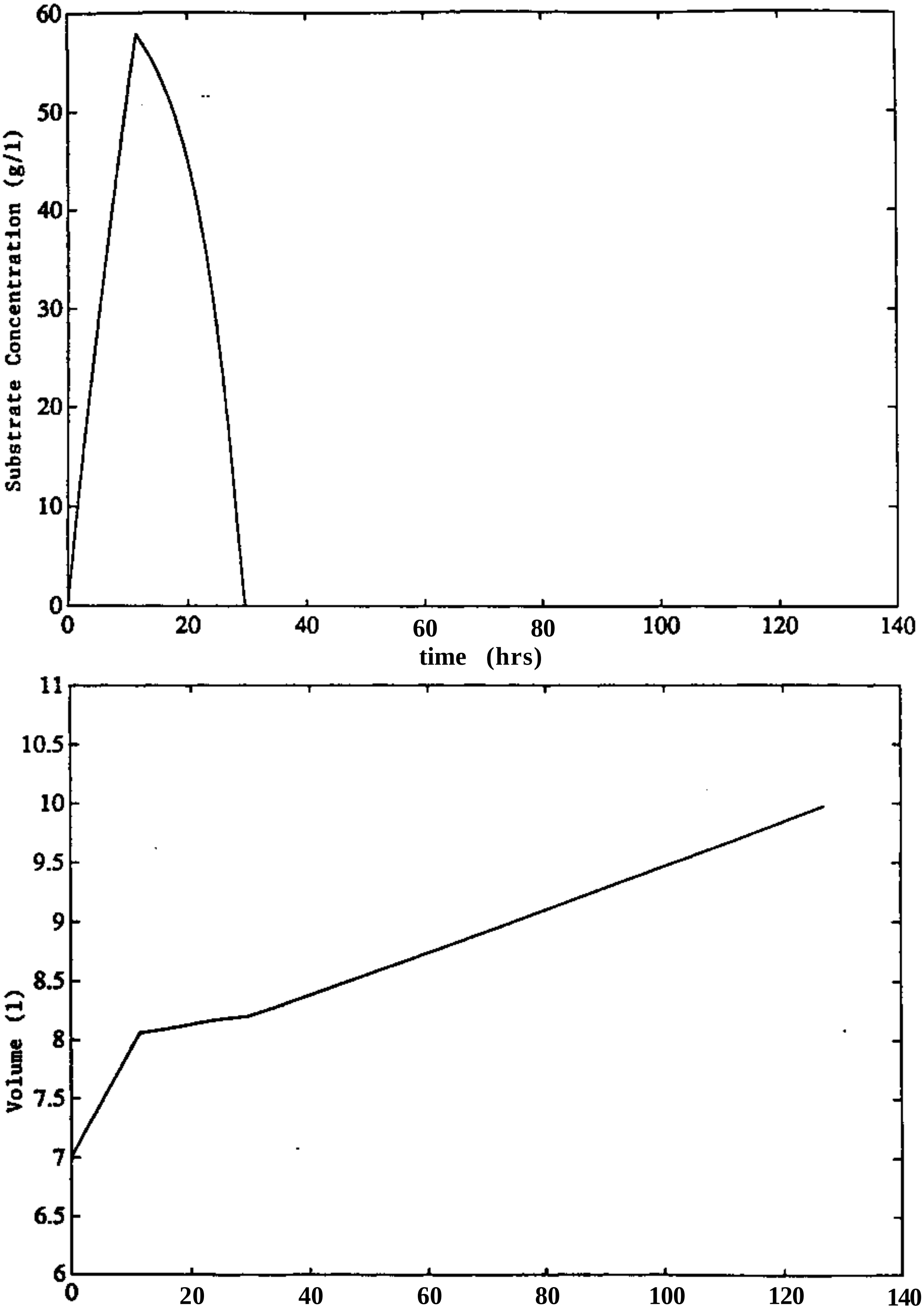


Figure 4-7: Substrate and Volume Profiles for Case I

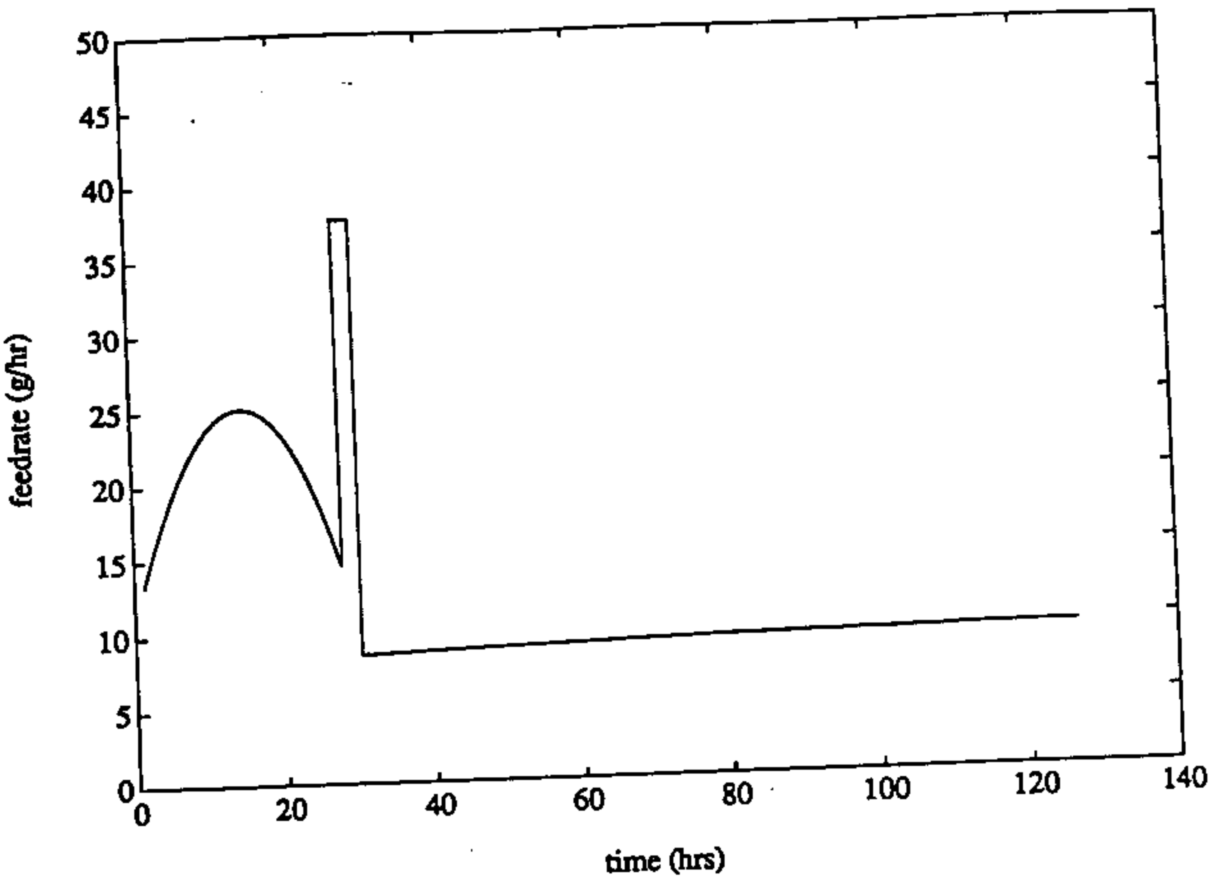


Figure 4-8: Control Profile for Case II

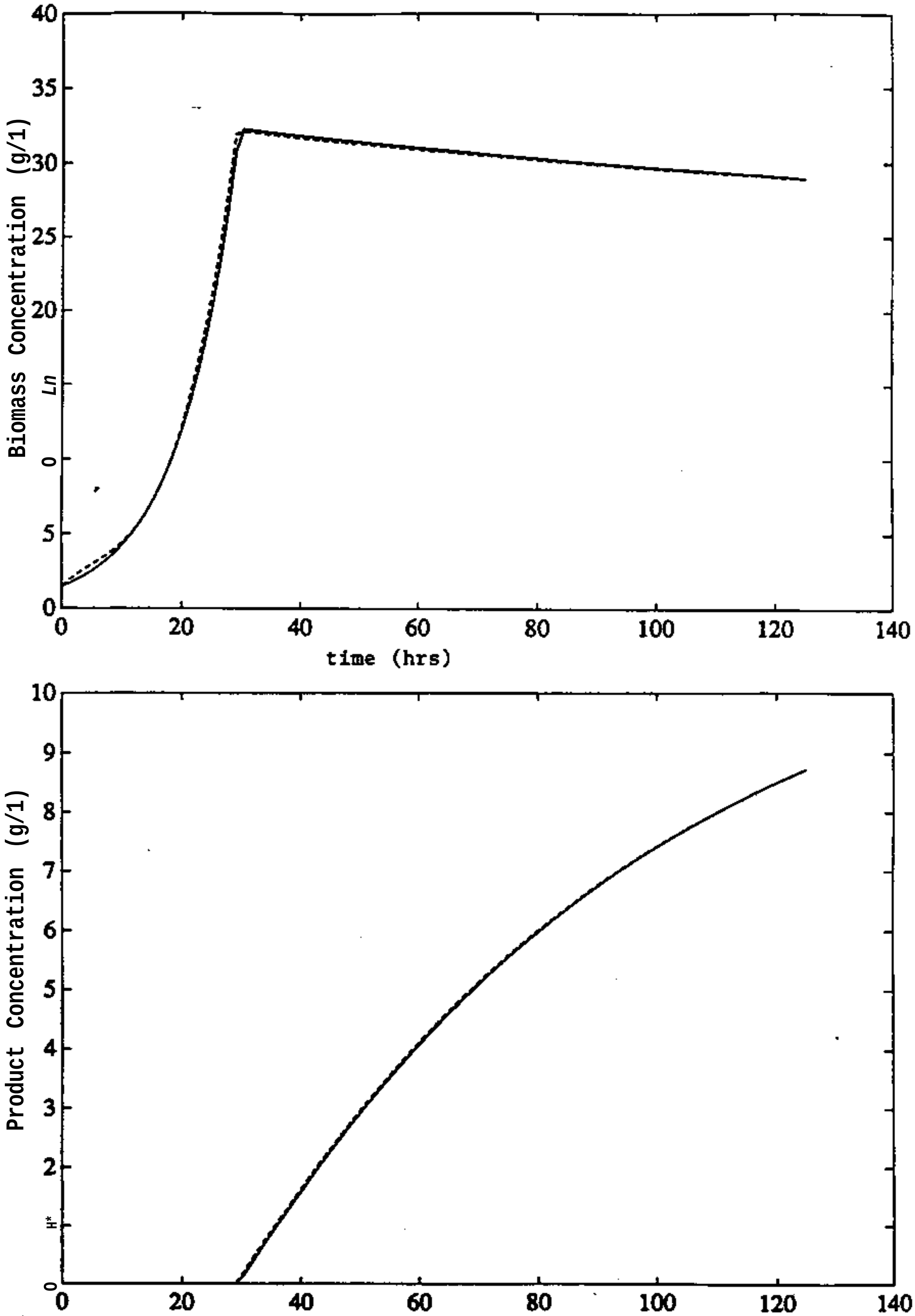


Figure 4-9: Biomass and Product Profiles for Case II

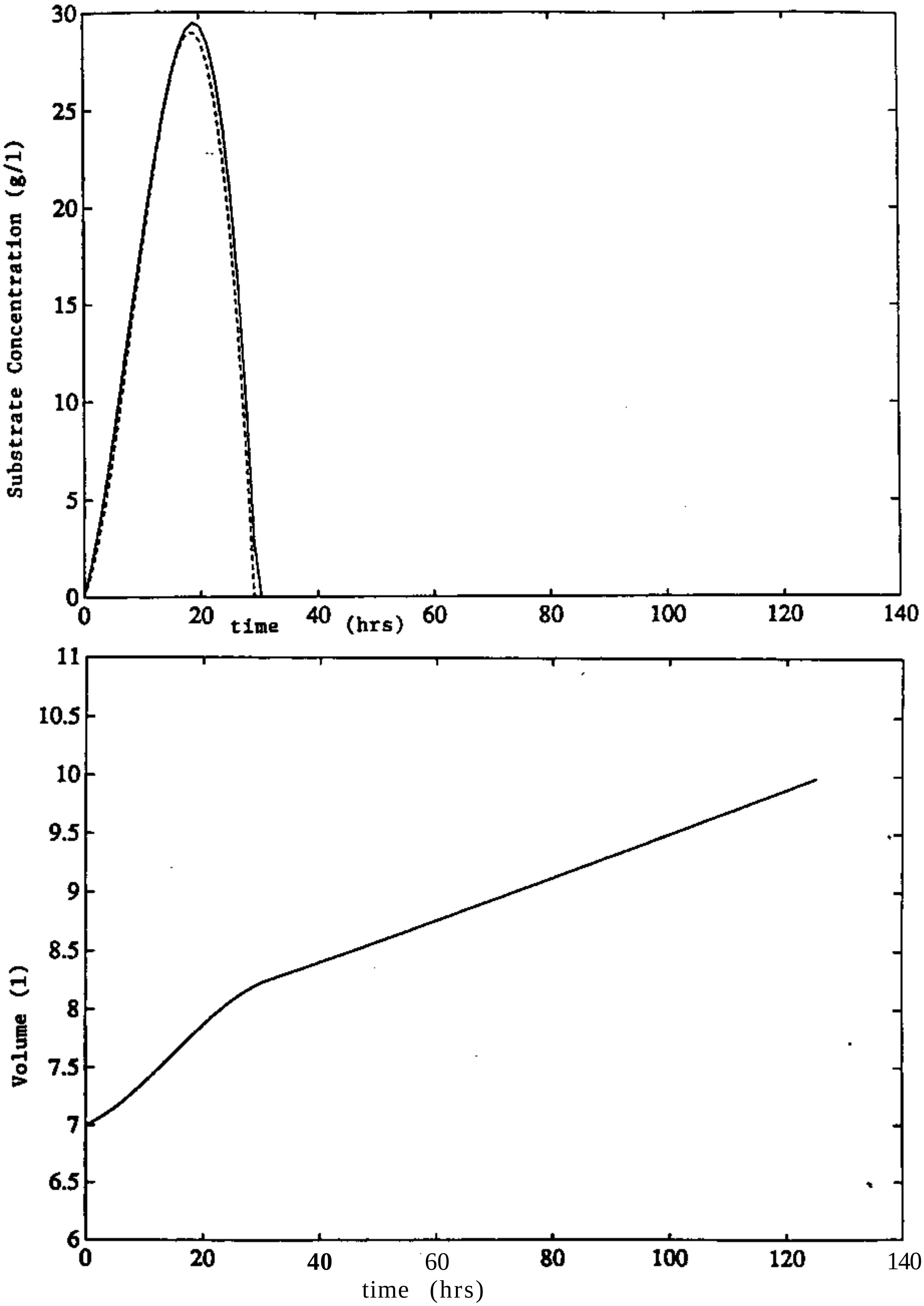


Figure 4-10: Substrate and Volume Profiles for Case II

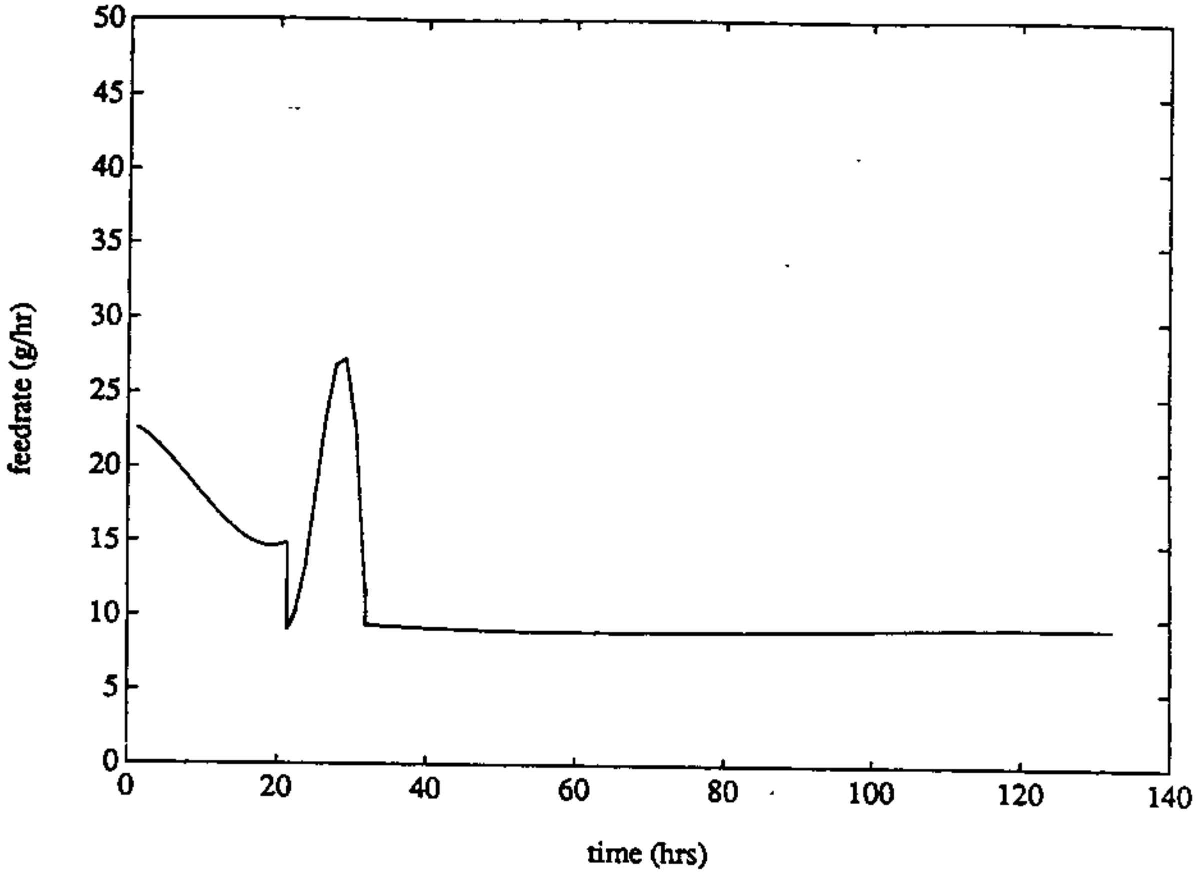


Figure 4-11: Control Profile for Case III

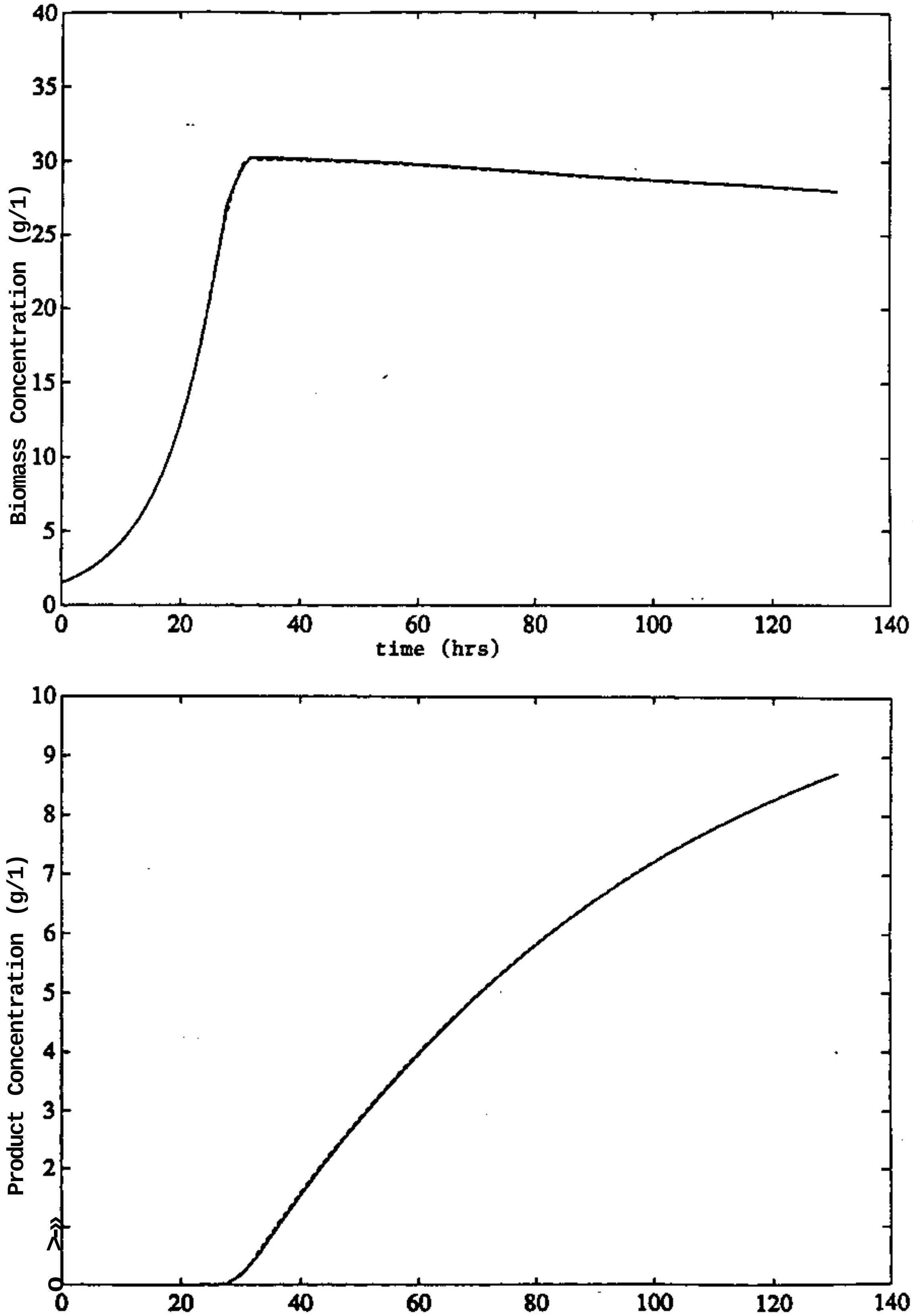


Figure 4-12: Biomass and Product Profiles for Case III

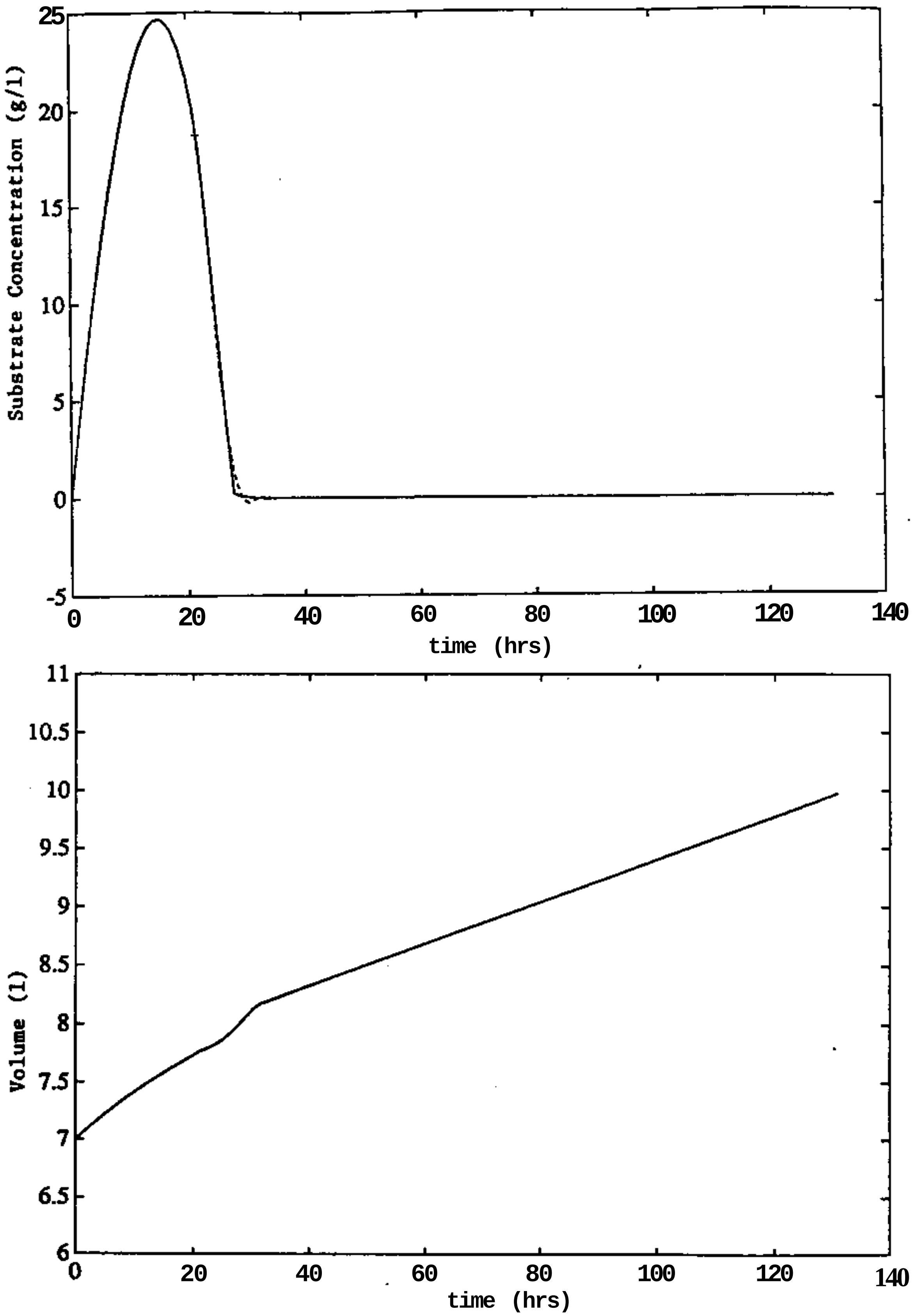


Figure 4-13: Substrate and Volume Profiles for Case III

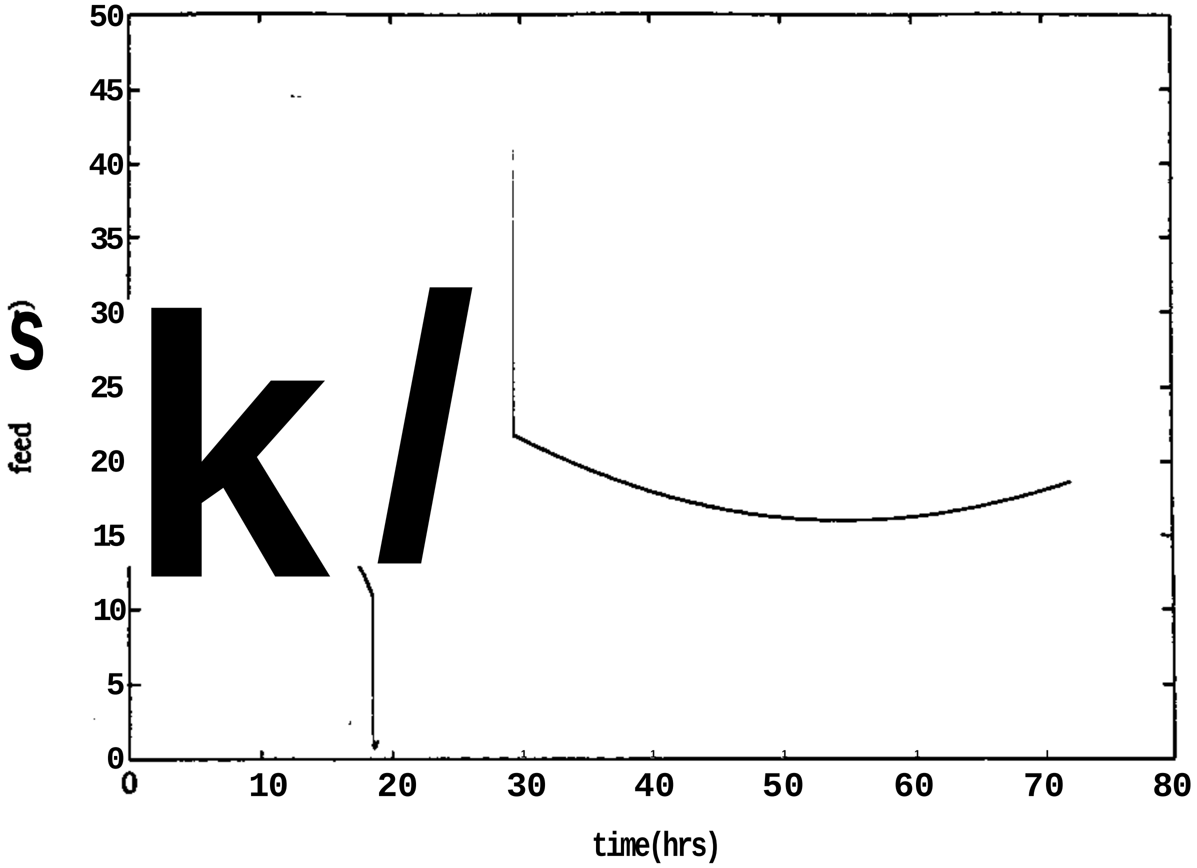


Figure 4-14: Control Profile for Case IV

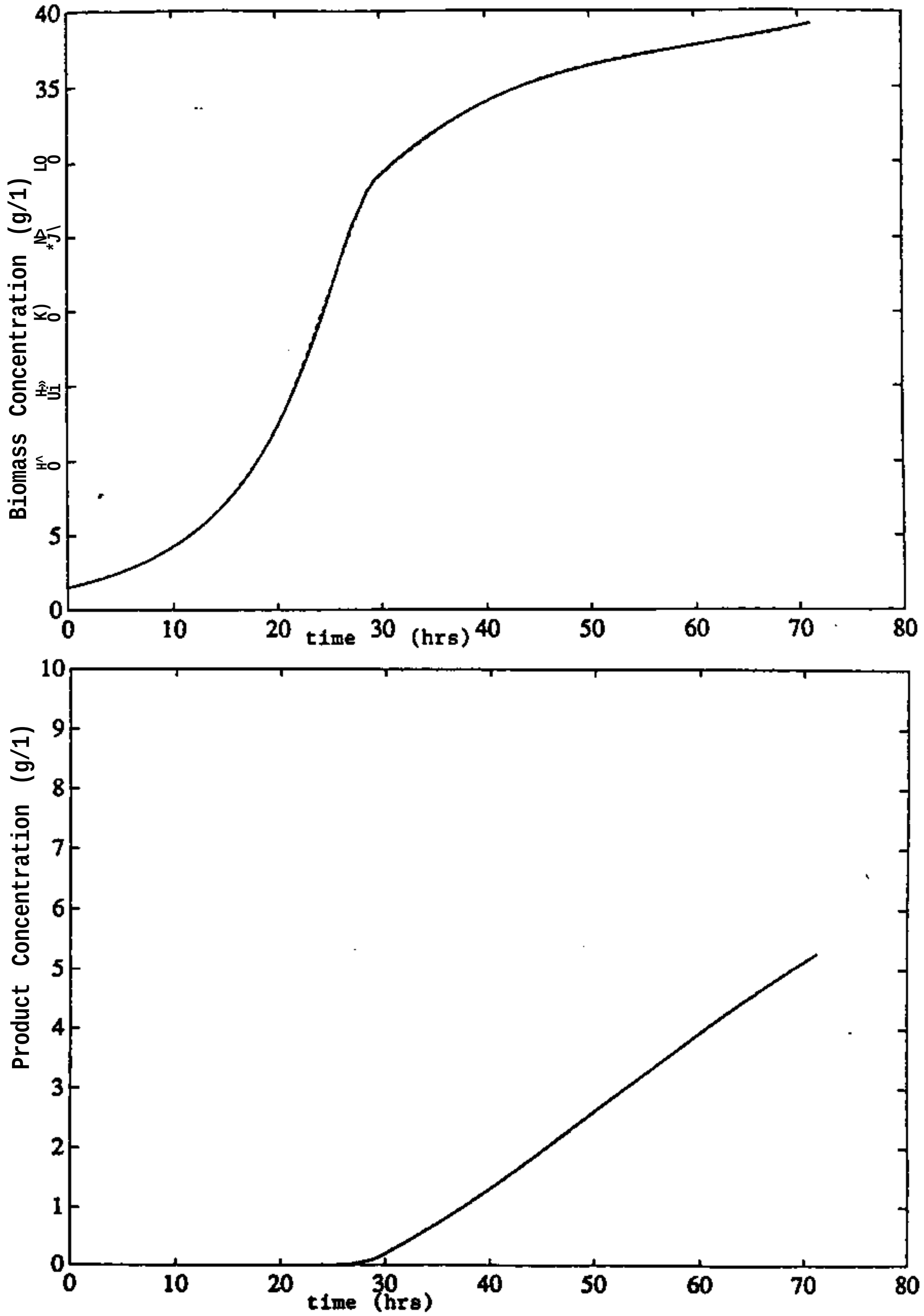
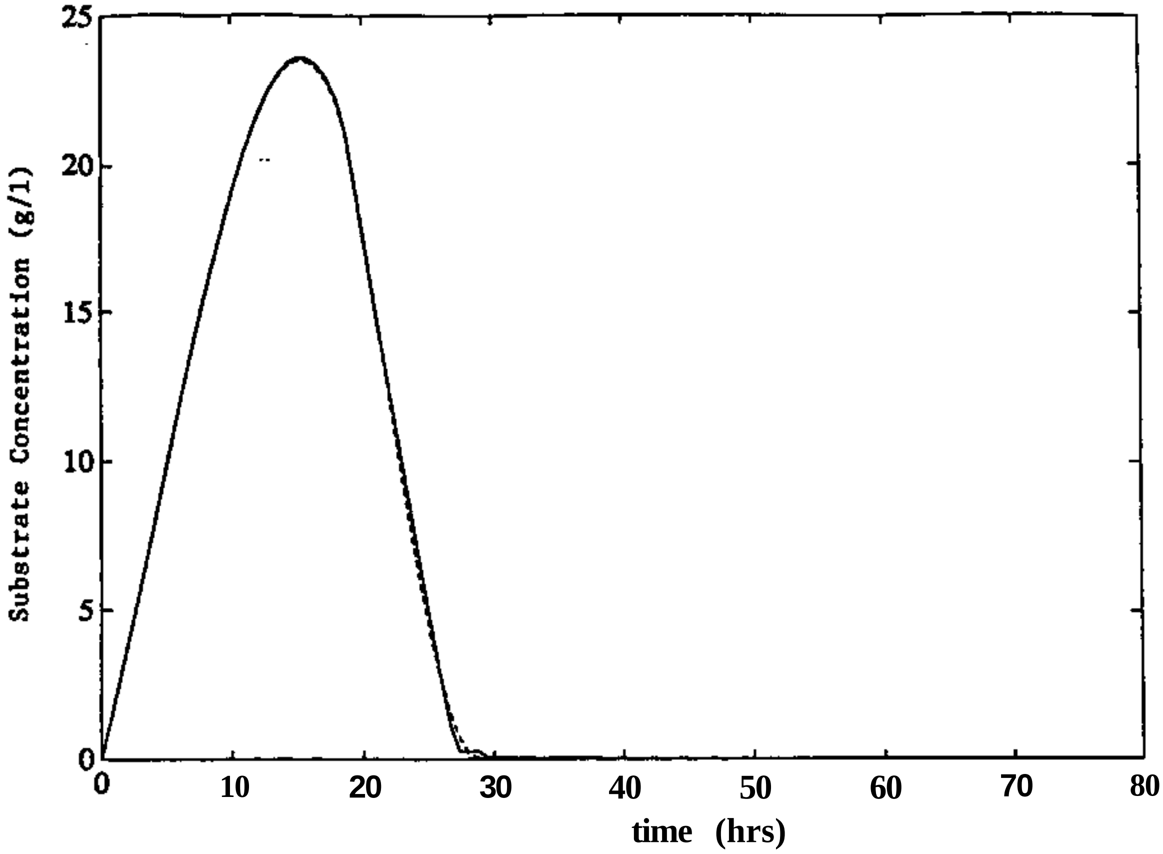


Figure 4-15: Biomass and Product Profiles for Case IV



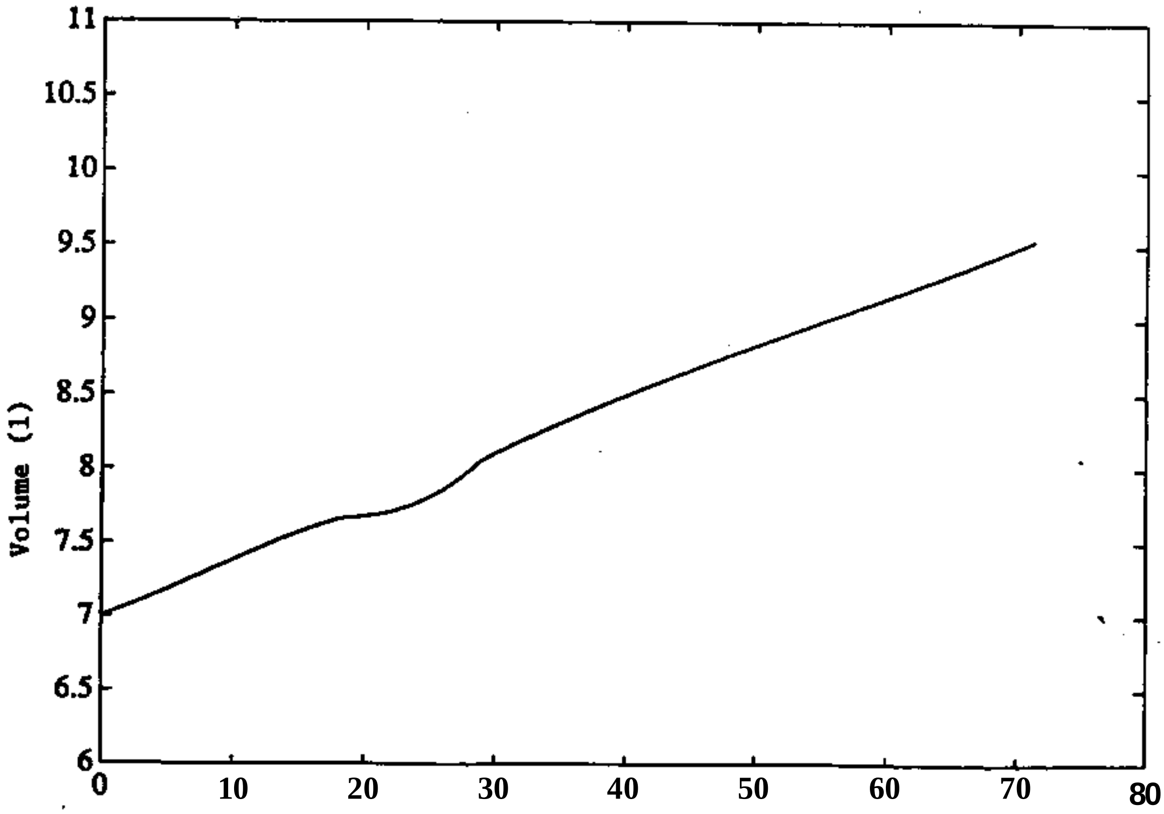


Figure 4-16: Substrate and Volume Profiles for Case IV

4.3. Comparison of Computation Aspects

The approach presented here for solving discontinuous control problems offers a number of advantages over other methods. Alternate solution methods for problems of this type couple a DAE (differential-algebraic) solver to an optimization routine. The DAE solver integrates the model in an inner loop while the optimizer functions in an outer loop to perform the optimization. This approach can be quite expensive, particularly if a gradient based optimizer (e.g., SQP) is used. In this case, the required gradients will likely be obtained through perturbation of the model, or through sensitivity equations. On the other hand, Lim et al. (1986) applied variational conditions to (BFP) and had to solve a two point boundary value problem. Here they repeatedly used a shooting method coupled with LSODE to optimize their profile, although this only had to be done over the final feeding portion of the batch.

Just as important as the above difficulties, is the inability of DAE solvers to explicitly handle *inequalities,* such as profile bounds. In the work of Modak et al. (1986) and Lim et al. (1986), the bounds on the control profile were handled implicitly, by knowing the portions of reaction time over which these bounds were active. The condition on the tank volume could, however, be considered an upper bound, but this was also handled implicitly by setting the flowrate to zero when the tank was full. For handling inequality constraints with an ODE (or DAE) solver, only one other option exists, that of Sargent and Sullivan (1977), However, even this approach does not overcome the problem of potentially expensive solutions.

Our approach, on the other hand, formulates a discrete version of (BFP) as an NLP, and the above difficulties are not encountered. However, two important considerations with our approach involve the choices of K and NE and solution of the resulting large-scale nonlinear program. The first problem is currently handled by trial and error solution. The second problem can be remedied by using large-scale SQP decomposition methods (see Vasantharajan and Biegler (1987), Locke, Edahl and Westerberg (1983)). These problems form the basis for future research.

5. Conclusions

This paper presents a general method for solving optimal control problems that have discontinuous profiles. The approach uses orthogonal collocation on finite elements to discretize the differential equation model, and Lagrange polynomials to construct approximations to the continuous profiles. The resulting algebraic modelling equations are written as constraints in a Nonlinear Program with polynomial coefficients becoming decision variables. The finite element knot locations are also written as variables in the NLP so that points of discontinuity can be found.

Two important theoretical properties have been addressed. First, approximation properties were considered through an equivalence between orthogonal collocation on finite elements and a fully implicit Runge-Kutta integration scheme that uses Gaussian points. As a result, desirable stability and high order accuracy properties, normally associated with numerical integration schemes, also hold for the orthogonal collocation method used here. The topic of accuracy of optimal control profiles was also addressed. Here the NLP approach outlined above was shown to be equivalent to solving discrete approximations of the variational necessary conditions. Specifically, equivalence was established between the Kuhn-Tucker conditions of the NLP and the discretized necessary variational conditions.

Applicability of the method has been demonstrated by solving a well-known fed-batch fermenter problem. Solutions slightly better than published, analytically-based results were obtained although nonunique solutions also are found due to insensitivity of the objective function to the optimal control profile. This insensitivity leads to slight variations in the optimal values of the points of control profile discontinuity, final batch time and the optimal state and control profile shapes. This occurred even though objective function values varied by less than 0.05%. Very good accuracy of the model approximations resulted and these clearly demonstrate the effectiveness of using orthogonal collocation and Lagrange polynomials. In addition the flexibility of the NLP approach was demonstrated by adding state variable inequality constraints on the substrate profile. This was easily done by changing the bounds on the polynomial coefficients in the NLP. With other methods that use a variational calculus approach, imposition of a state profile inequality constraint can only be done implicitly and is much more difficult

References

Ascher, U. and G. Bader, "Stability of Collocation at Gaussian Points/[1] *SIAM J. Numer. Anal.*, 23(2), pp. 412-422, (1986).

Bajpai, R. K. and M. Reuss, "A Mechanistic Model for Penicillin Production/[1] *J. Chem. Tech. Biotechnol.*, 30, pp. 332-344, (1980).

Bajpai, R. K. and M. Reuss, "Evaluation of Feeding Strategies in Carbon-Regulated Secondary Metabolite Production through Mathematical Modelling," *Biotechnol. Bioeng.*, 23, pp. 714-738, (1981).

Biegler, L. T. and J. E. Cuthrell, "Improved Infeasible Path Optimization for Sequential Modular Simulators-II: The Optimization Algorithm," *Comp. and ChE.*, 9(3), pp. 257-267, (1985).

Bryson, A. E. and Y. -C. Ho, Applied Optimal Control, Hemisphere Publishing, Washington D.C., (1975).

Burrage, K. and J. C. Butcher, "Stability Criteria for Implicit Runge Kutta Methods," *SIAM J. Numer. Anal.*, 16(1) pp. 44-57, (1979).

Carnahan B., H. A. Luther and J. O. Wilkes, Applied Numerical Methods, Wiley, New York (1968).

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

Cuthrell, J. E and L T. Biegler, "On the Optimization of Differential-Algebraic Process Systems", *AlChE J.*, 33, 8, pp. 1257-1270 (1987)

Dahlquist, G., "Stability Questions for Some Numerical Methods for Ordinary Differential Equations," *Proc. Symp. Appl. Math.*, 15, p. 47, (1963).

Ehle, B. L, "High Order A-stable Methods for the Numerical Solution of Differential Equations," *BIT*, 8, p. 276, (1968).

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

Gardini, L., A. Servida, M. Morbidelli and S. Carra, "Use of Orthogonal Collocation on Finite Elements with Moving Boundaries For Fixed Bed Catalytic Reactor Simulation," *Comp. and ChE.*, 9(1), pp. 1-17, (1985).

Hindmarsh, A. C, "LSODE and LSODI, Two new initial value ordinary differential equation solvers/[1] ACM-SIGNUM Newsletter, 15(4), pp. 10-11, (1980).

Jacobsen, D. H., M. M. Lele, and J. L. Speyer, "New Necessary Conditions of Optimality for Control Problems with State-Variable Inequality Constraints/' *J. Math. Ana/, and Appl.*, 34. PP.239-266, (1971).

Kreindler, E., "Additional Necessary Conditions for Optimal Control with State-Variable Inequality Constraints," *JOT A*, 38(2), pp. 241-250, (1982).

Lim, H. C, Y. J. Tayeb, J. M. Modak and P. Bonte, "Computational Algorithms for Optimal Feed Rates for a Class of Fed-Batch Fermentation: Numerical Results for Penicillin and Cell Mass Production," *Biotech no I. Bioeng.*, 28, pp. 1408-1420, (1986).

Locke, M.H., R. Edahl and A.W. Westerberg, "An Improved Successive Quadratic Programming Optimization Algorithm for Engineering Design Problems," AIChE J, 29, 5, (1983)

Lynn, L. L., E. S. Parkin and R. L. Zahradnik, "Near-Optimal Control by Trajectory Approximation," *I&EC Fundamentals*, 9(1), pp. 58-63, (1970).

Lynn, L L and R. L. Zahradnik, "The Use of Orthogonal Polynomials in The Near-Optimal Control of Distributed Systems by Trajectory Approximation," *Int. J. Control*, 12(6), pp. 1079-1087, (1970).

Modak, J. M., personal communication, (1987).

Modak, J. M., H. C. Lim and Y. J. Tayeb, "General Characteristics of Optimal Feed Rate Profiles for Various Fed-Batch Fermentation Processes," *Biotechnol. Bioeng.*, 28, pp. 1394-1407, (1986).

Neuman, C P. and A. Sen, "A Suboptimal Control Algorithm for Constrained Problems Using Cubic Splines," *Automatic^*, 9, pp. 601-613, (1973a).

Neuman, C P. and A. Sen, "Weighted Residual Methods and the Suboptimal Control of Distributed Parameter Systems," *Int. J. Control*, 18(6), pp. 1291-1301, (1973b).

Oh, S. H. and R. Luus, "Use of Orthogonal Collocation Method in Optimal Control Problems," *Int. J. Control*, 26(5), pp. 657-673, (1977).

Reddien, G. W., "Collocation at Gauss Points as a Discretization in Optimal Control," *SIAM J. Control and Optimization*, 17(2), pp. 298-306, (1979).

Sargent, R. W. H. and G. R. Sullivan, "The Development of an Efficient Optimal Control

Package/[1] *Proceedings of the 8th IFIP Conference on Optimization Techniques Pt* Z, (1977).

Tayeb, Y. J. and H. C Lim, "Optimal Glucose Feed Rates for Fed-Batch Penicillin Fermentation: An Efficient Algorithm and Computational Results/* *Annals New York Academy of Sciences,4*69, pp. 382-403, (1986).

Tsang, T. H., D. M. Himxnelblau and T. F. Edgar, "Optimal Control via Collocation and Non-linear Programming," *Int. J. Control,* 21(5), pp. 763-768, (1975).

Vasantharajan, S. and L.T. Blegler, "Large-Scale Decomposition for Successive Quadratic Programming," submitted for publication (1987)

Villadsen J. and M. L. Michelsen, Solution of Differential Equation Models by Polynomial Approximation, Prentice-Hall, (1978).

Wong, K. T. knd R. Luus, 'Time Suboptimal Feedback Control of Systems Described by Linear Parabolic Partial Differential Equations," *Optimal Control Applications & Methods,* 3, pp. 177-185, (1982).