Brazilian Journal of Chemical Engineering

ISSN 0104-6632
Printed in Brazil
www.abeq.org.br/bjche

Vol. 24, No. 03, pp. 445 - 452, July - September, 2007

A NOVEL HYBRID OPTIMIZATION ALGORITHM FOR DIFFERENTIAL-ALGEBRAIC CONTROL PROBLEMS

F. S. Lobato[1], L. C. Oliveira-Lopes[2] and V.V. Murata[3*]

School of Chemical Engineering, Federal University of Uberlândia
Campus Santa Mônica, 38400-902, Uberlândia - MG, Brazil.
E-mail: franpi22@yahoo.com.br[1], E-mail: lcol@ufu.br[2],
E-mail: valeria@ufu.br[3]

(Received: January 5, 2006 ; Accepted: June 6, 2007)

Abstract - Dynamic optimization problems can be numerically solved by direct, indirect and Hamilton-Jacobi-Bellman methods. In this paper, the differential-algebraic approach is incorporated into a hybrid method, extending the concepts of structural and differential indexes, consistent initialization analysis, index reduction and dynamic degrees of freedom to the optimal control problem. The resultant differential-algebraic optimal control problem is solved in the following steps: transformation of the original problem into a standard nonlinear programming problem that provides control and state variables, switching time estimates and costate variables profiles with the DIRCOL code; definition of the switching function and the automatically generated sequence of index-1 differential-algebraic boundary value problems from Pontryagin's minimum principle by using the developed Otima code; and finally, application of the COLDAE code with the results of the direct method as an initial guess. The proposed hybrid method is illustrated with a pressure-constrained batch reactor optimization problem associated with the slack variable method.
Keywords: Optimal control; Differential - algebraic equations: Hybrid method.

INTRODUCTION

A dynamic optimization problem, also known as optimal control problem (OCP), consists in determination of the control variable profiles that maximize or minimize a measure of performance. The significant increase in its application in industry over the past decade has been mainly due to the high popularity of dynamic simulation tools associated with a competitive global market, in which environmental constraints and demanding market specifications require a continuous optimization of process operations. Dynamic optimization enables an automatic decision-making procedure and as it gets established as a useful and trustworthy technology, it will foment other applications, such as the addressing of hard-constrained problems, the synthesis of chemical reactor networks, the uncertainty description in multiple period problems and the development of tools such as automatic differentiation (Biegler et al., 2002).

The methods for solution of OCP are divided in direct (Cervantes and Biegler, 1999), indirect (Pontryagin et al., 1962; Bryson and Ho, 1975; Ray, 1981) and dynamic programming (Bellman, 1957) methods.

The direct approach uses control parameterization (the sequential method) or state

*To whom correspondence should be addressed

and control parameterizations (the simultaneous method), transforming the original problem into a finite dimensional optimization problem. By all means, the implementation of direct methods is simpler because it does not require the generation of the costate or adjoint equations, which, at the very least, will duplicate the dimensions of the set of differential-algebraic equations (DAE) in the indirect method. On the other hand, the solution of nonlinear programming problems (NLP) of great dimension or the attainment of the gradients of the objective function in the sequential method is not trivial (Feehery, 1998). The direct method has been preferentially used in the last several years (Stryk and Bulirsch, 1992; Feehery, 1998; Cervantes and Biegler, 1999; Biegler et al., 2002; Huang et al., 2002). In spite of its accuracy being lower than that of the indirect methods, this preference is likely due to the extended convergence properties of NLP algorithms (Bulirsch et al., 1991a, 1993).

The indirect strategy for solution of OCP is based on variational principles. These conditions, from Pontryagin's maximum principle (Bryson and Ho, 1975), generate a set of Euler-Lagrange equations, which are boundary value problems (BVP), inherently formed by differential-algebraic equations regardless of whether or not the problem is constrained. Some difficulties in the OCP solution must be highlighted: the existence of end-point conditions or region constraints gives rise to multipliers and associated complementary conditions that significantly increase the difficulty of solving the BVP by the indirect method; the existence of constraints in the state variables and application of the slack variables method may originate differential-algebraic optimal control problems (DAOCP) with higher indexes, regardless of the constraint activation status, even in problems where the number of inequality constraints is equal to the number of control variables and the Lagrange multipliers may be very sensitive to the initial conditions. The BVP can be solved by shooting methods, collocation on finite elements or finite difference schemes and may present convergence problems due to the difficulty in supplying adequate initial estimates for the costate variables profiles. The direct methods do not have this problem, but they may generate low-precision and suboptimal solutions. Bulirsch et al. (1991a-b, 1993) and Koslik and Breitner (1997) used homotopic techniques and an initial simplified problem to overcome the difficulties originating from bang-bang arcs and activation and deactivation of inequality constraints. In general, the inequality constraints are neglected and the resulting simplified problem is solved by direct methods. In sequence, this solution is used as initial estimates for the indirect method.

Most recently, the combination of direct and indirect methods as an alternative to retain the best of each characteristic was proposed, resulting in the so-called hybrid methods (Bulirsch et al., 1991a-b, 1993; Koslik and Breitner, 1997). However, to date results are largely case-dependent and as yet inconclusive regarding the global efficiency of the approach.

The significant advances in mathematical symbolic tools and the differential-algebraic approach to modeling and process simulation are also remarkable. Application of the differential-algebraic approach to optimal control problems necessarily requires the extension of well-established concepts in the field of dynamic simulation, such as number of degrees of freedom, differential index, structural index, consistency of initial conditions, drift-off of constraints, index reduction effects and so forth. The number of dynamic degrees of freedom is associated with the number of initial conditions of the original problem and with the number of boundary value conditions of the augmented problem (Feehery, 1998). These extensions must also consider the numerical methods for integration of differential-algebraic boundary problems, which have not reached the same level of development as the methods for initial value problems; the solution of singular arc optimization problems, which will occur in affine control problems and relate to intrinsic higher index equations; the fact that index reduction might also be mandatory for the establishment of the correct number of boundary conditions associated with the set of DAE and the index fluctuation along the optimal trajectory due to the activation and deactivation of inequality constraints on the state variables.

The main goal of this contribution is to introduce a general procedure to systematize the solution of differential-algebraic optimal control problems (DAOCP) using a hybrid method, which combines the best characteristics of the direct and indirect methods in a differential-algebraic framework, and is organized as follows. Section 2 briefly presents the general concepts of dynamic optimization, differential-algebraic approach to formulating an optimal control problem, numerical methods for

DAE solution and DAE characterization as a necessary background for introducing the proposed hybrid algorithm, applied to DAOCP of fluctuating index. In Section 3, a case study of its application to a constrained batch reactor system is presented. Finally, the conclusions are drawn in Section 4.

PROPOSED HYBRID ALGORITHM FOR DIFFERENTIAL-ALGEBRAIC OPTIMAL CONTROL PROBLEMS

Characterization of DAOCP can be made by extending the concepts used in the algebraic-differential approach to simulation problems. This characterization can be implemented using structural tools, such as the ALGO and PALGO codes (Unger et al., 1995), which can supply information beyond the index, the number of degrees of freedom and the variable-equation relationship to facilitate manipulation and analysis of system consistency. Another important aspect of this characterization is that, in contrast to what occurs in dynamic problems where consistent initial values are known or can be determined, some components of the initial state are free or depend on the control. Therefore, consistency of the initialization must be guaranteed during numerical solution of the problem of optimal control (Gerdts, 2001).

The proposed methodology splits the original problem into index-1 phases, through determination of events based on the constraint activation and deactivation obtained by the direct approach. The necessary characterization of each phase is addressed by using structural concepts and a near optimal solution for the events and costate variables given by the direct method are used as initial guesses for the two-point boundary value problem of index-1 solved by the indirect method. The main steps of the proposed algorithm are presented in Figure 1 and are described in sequence:

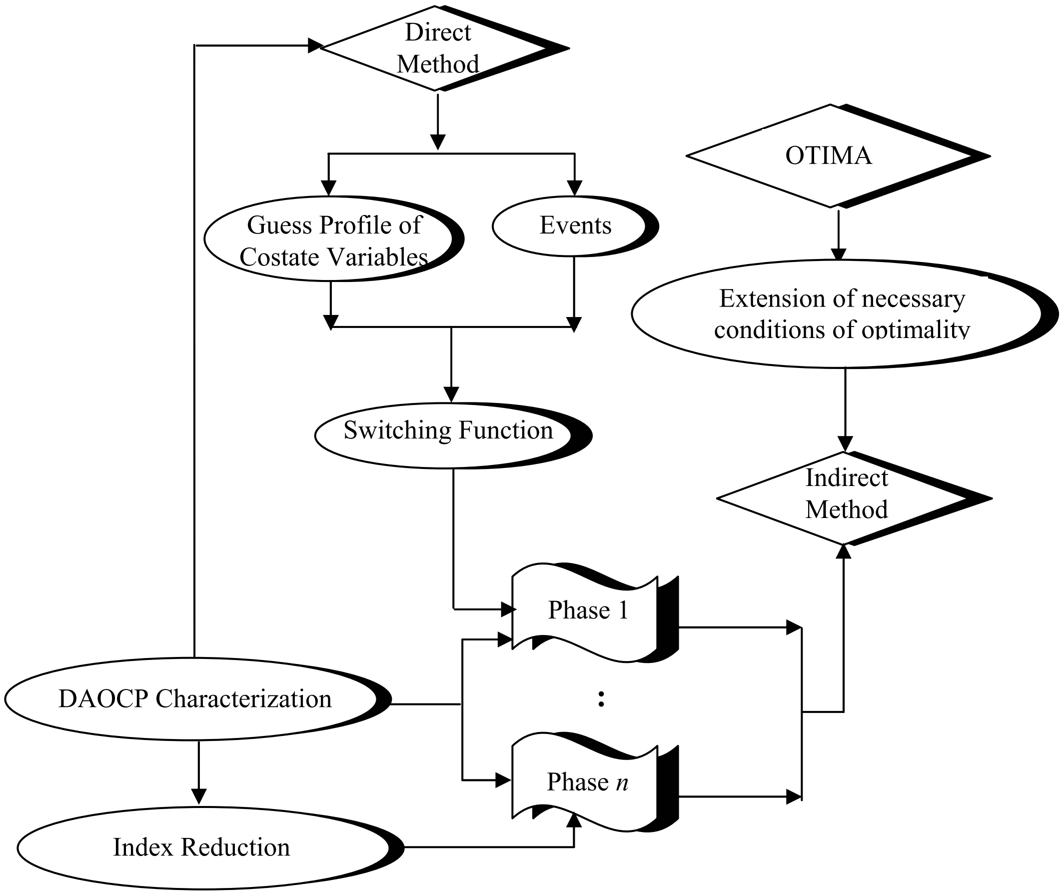


Figure 1: Hybrid algorithm for differential-algebraic optimal control problems

- Step 1: Structural characterization of the original problem through ALGO and PALGO codes (Unger et al., 1995), for the structural indexes and the number of degrees of freedom;
- Step 2: Index reduction to one for higher index original problems, according to the information given by structural codes;
- Step 3: Numerical solution of the original DAOCP using the direct approach, with the DIRCOL code (Stryk, 1999), which provides identification of the events, estimation for costate variable profiles and activations and deactivations constraints and calculation of the control moves in each phase;
- Step 4: Automatic generation of Euler-Lagrange equations for the augmented system based on an extended version of the OTIMA code implemented in the Maple® environment (Gomes, 2000; Lobato, 2004) according to the differential-algebraic approach (Feehery, 1998);
- Step 5: Structural characterization of the augmented system generated in Step 4;
- Step 6: Index reduction of the higher index boundary value problem defined by each phase;
- Step 7: Solution of the resulting DAOCP with the indirect method approach through the COLDAE code (Ascher and Spiteri, 1994), using the results from Step 3 as the initial guess to assure convergence, with once the consistency of initial conditions in each phase is guaranteed.

The direct method was implemented using both control and state variable parameterization (Stryk, 1999) and the indirect method was based on a collocation procedure (Ascher and Spiteri, 1994). These codes were chosen due to their flexibility and the desired differential-algebraic characteristics (Lobato, 2004; Santos et al., 2005). The proposed algorithm may drive the solution to satisfy the constraints and despite the consequently higher computational demand, it might enforce the solution for a better value for the objective function. The next section illustrates the ability of the algorithm in solving a benchmark constrained batch reactor problem.

APPLICATION: A BATCH REACTOR CASE STUDY

This section presents the optimal solution for an isothermal batch reactor system operating in gas phase (Feehery, 1998; Huang et al., 2002). The dynamic optimization of this constrained problem is given by:

$$\min\ x_3(t_f) \tag{1}$$

$$\dot{x}_1 = -k_1 x_1 + k_2 x_2^2 + u/V - k_3 x_1 x_2 \tag{2}$$

$$\dot{x}_2 = k_1 x_1 - k_2 x_2^2 - k_3 x_1 x_2 \tag{3}$$

$$\dot{x}_3 = k_3 x_1 x_2 \tag{4}$$

$$N = V(x_1 + x_2 + x_3) \tag{5}$$

$$PV = NRT \tag{6}$$

$$P \le P_{max} \tag{7}$$

$$u_{min} \le u \le u_{max} \tag{8}$$

$$x(0) = [100, 0, 0]^T \tag{9}$$

The variables x_1, x_2 and x_3 represent the species concentration of A, B and D (mol/m^3) respectively; P is the pressure of the reactor (Pa); N is the total number of moles (mol); V is the volume of the reactor (1 m^3); R is the constant of gases (8.314 J/(mol K)); u is the pure feed flow rate of A (mol/hr); k_1, k_2 and k_3 are the kinetics reaction constants (0.8 hr^{-1}, 0.02 m^3/(mol hr), 0.003 m^3/(mol hr), respectively) and T is the operating temperature (400 K). The system is constrained to operate subject to the following limits: P_{max} = 340000 Pa, u_{min} = 0.0 mol/hr and u_{max} = 8.5 mol/hr.

Feehery (1998) presented a solution to this problem using finite elements to approximate the control profile, addressing the fact that the control is not independent during constraint activation, and through the development of an algorithm that is a significant improvement in evaluation of the sensitivity equations. In Huang et al. (2002) this problem was dealt with through a decomposition strategy, based on a combination of the advantages of the simultaneous and sequential approaches, which uses the following criterion of partition: a new system (Sis1) contains the state variables that appear in the inequality constraint P and the remaining variables x_1, x_2, x_3 and N migrate to another system (Sis2). The general idea is to use the Sis2 system to attain the equations of sensitivity necessary for the solution of system Sis1.

In the following, the costate equations (Equations 10-17); the stationary condition (Equation 18); and the boundary conditions for the costate variable λ, defined at t_f generated in Step 4 of the proposed algorithm, are shown as an output of the OTIMA code (Lobato, 2004):

$$\dot{\lambda}_1 = (k_1 + k_3 x_2)\lambda_1 + (-k_1 + k_3 x_2)\lambda_2 - -k_3 x_2 \lambda_3 - V\lambda_4 \quad (10)$$

$$\dot{\lambda}_2 = (-2k_2 x_2 + k_3 x_1)\lambda_1 + +(2k_2 x_2 + k_3 x_1)\lambda_2 - k_3 x_1 \lambda_3 - V\lambda_4 \quad (11)$$

$$\dot{\lambda}_3 = -V\lambda_4 \quad (12)$$

$$\lambda_4 = RT\lambda_5 \quad (13)$$

$$\lambda_5 V = -\lambda_6 \quad (14)$$

$$\lambda_6 \delta_1 = 0 \quad (15)$$

$$\lambda_7 \delta_2 = 0 \quad (16)$$

$$\lambda_8 \delta_3 = 0 \quad (17)$$

$$-\lambda_1 / V - \lambda_7 + \lambda_8 = 0 \quad (18)$$

$$\lambda(t_f) = [0; 0; -1] \quad (19)$$

where δ is a vector of slack variables (Jacobson and Lele, 1969) associated with the constraint inequalities given by Equations (7) and (8).

The results of the application of structural ALGO and PALGO codes to the extended system formed by Equations (1) through (18) indicate a structural index-2 system. It must be highlighted that for problems with inequality constraints, the given results refer to the problem with active constraints, which means that since the index fluctuates, this is the largest value in all phases of the problem. The results obtained from DIRCOL show that this problem can be divided into two consecutive phases, defined for the constrained activation in the state variable P, with index-1 and index-2, respectively.

Furthermore, the control variable can be defined as $u=u_{max}$ for the first phase ($0<t<t_s=0.4740$) and $u=k_3Vx_1x_2$ ($t_s=0.4740<t<t_f=2$) for the second one due to the constraint activation and differentiation of Equation (7) with respect to time.

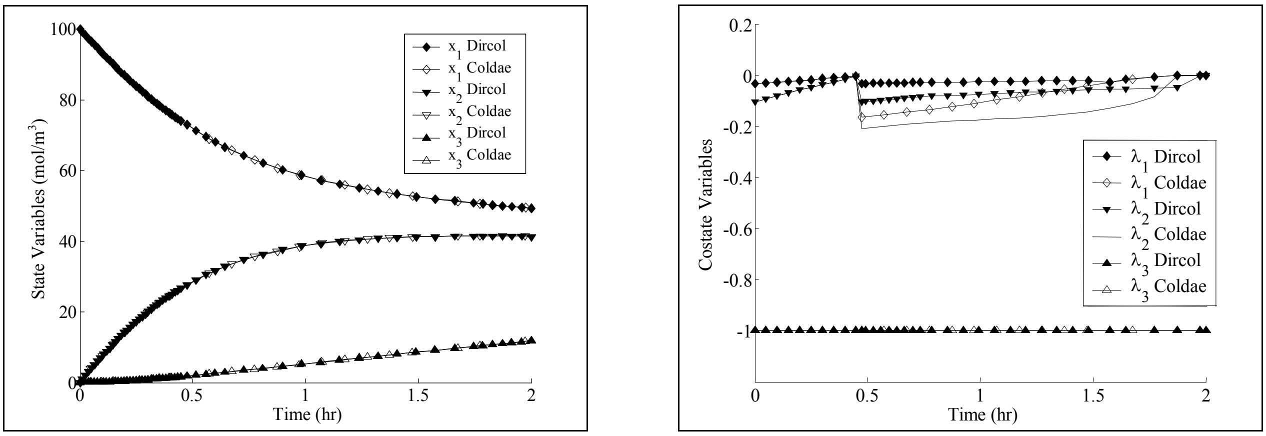
The state variable profiles obtained in Steps 3 and 7 of the proposed methodology are presented in Figures (2a), (3b) and (4a) and are comparable with previous results from the literature (Feehery, 1998; Huang, 2002). However, the profiles for the costate variables supplied by DIRCOL, given in Figure (2b), do not satisfy the boundary condition defined at $t=t_f$. This fact is not observed when analyzing the results obtained with the COLDAE code, which meets the conditions in a much better fashion, as shown at Table 1.

Figure (4b) shows the profiles for the slack variables. In the first phase, as u assumes its maximum value, δ_2 and δ_3 acquire the maximum and minimum values, respectively. In the second phase, as δ_1 vanishes due to the constraint activation, δ_2 and δ_3 assume intermediary values.

The value of the objective function of this work and those reported in the literature are provided in Table 2, where NC and NE are the number of collocation points and number of elements used in the numerical solution of the problem by the respective authors.

Table 1: Fulfillment of the costate variables at $t=t_f$.

Code	$\lambda_1(t_f)$	$\lambda_2(t_f)$	$\lambda_3(t_f)$
DIRCOL	$6.77\ 10^{-3}$	$2.53\ 10^{-3}$	1.00058
COLDAE	$8.56\ 10^{-8}$	$1.13\ 10^{-8}$	1.00000

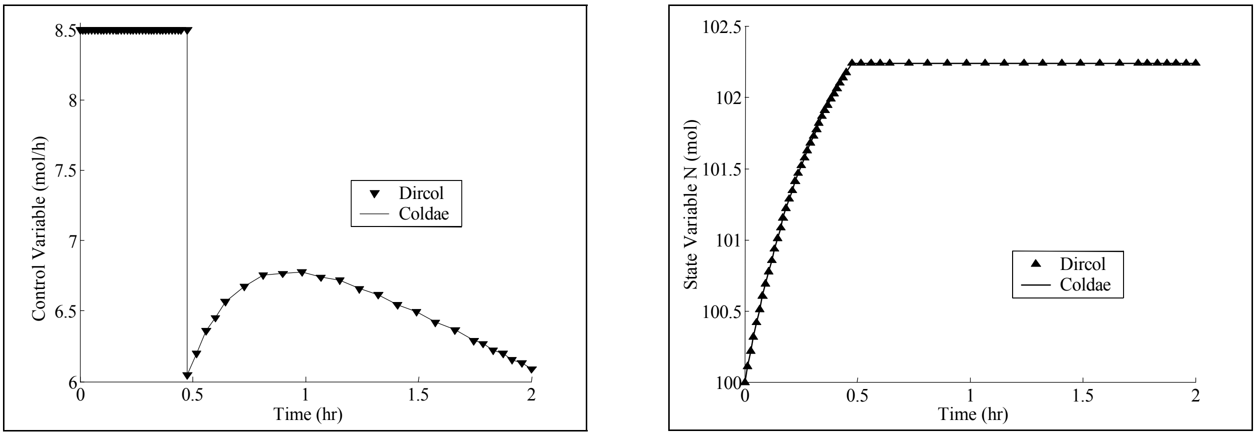


(a) (b)

Figure 2: (a) State variable profiles, (b) Costate variable profiles obtained with the DIRCOL and COLDAE codes

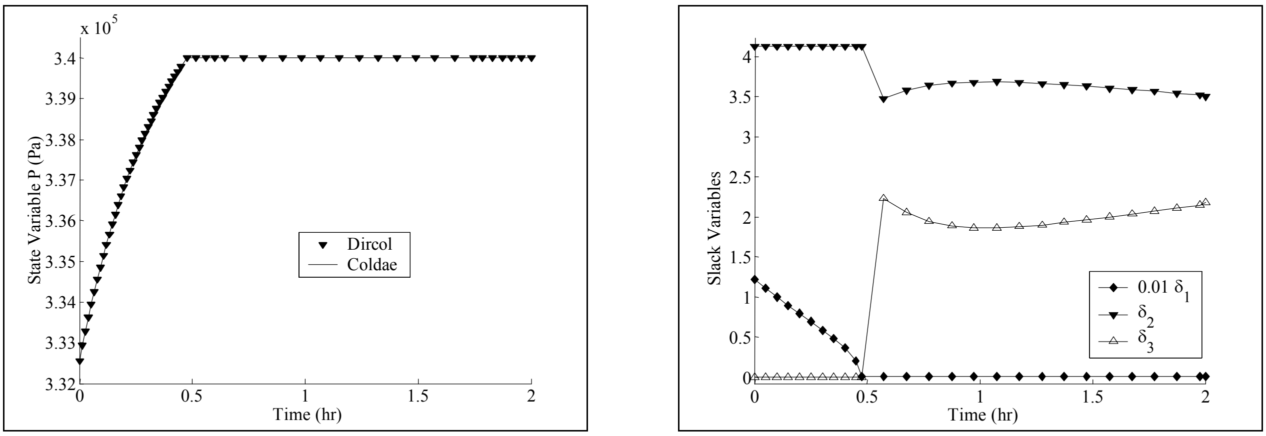
Table 2: Objective function value for constrained batch reactor optimization model.

Reference	Objective Function	Tolerance	NC/NE
Feehery (1998)	11.7284	10^{-7}	-/2
Huang et al. (2002)	11.7446	10^{-6}	11/1
This work	11.7255	10^{-7}	6/5



(a) (b)

Figure 3: (a) Control variable profile, (b) State variable N profile obtained with the DIRCOL and COLDAE codes.



(a) (b)

Figure 4: (a) State variable P profile, (b) Slack variable profiles obtained with the DIRCOL and COLDAE codes.

CONCLUSIONS

The proposed algorithm eliminates the index fluctuations of problems with state constraint inequalities or control affine optimization problems at the expense of requiring additional effort for the application of optimality conditions extended to DAE and index reduction in each identified phase. However, these additional tasks are greatly facilitated by the use of the symbolic OTIMA code and of information supplied by the ALGO and PALGO structural analysis codes.

The results for the objective function in the hybrid methodology were systematically better than the ones obtained by the direct method approach. This was also true to other benchmark cases presented elsewhere (Lobato, 2004). Although improvement of the solution may be considered negligible by many, it can be quite significant in cases where precision is a relevant factor, especially for guaranteeing the elimination of pseudo-optimal solutions, frequent in the study of discrete problems when using the direct method approach.

ACKNOWLEDGMENTS

The authors gratefully acknowledge the partial financial support of this work received from CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior), a Brazilian federal agency.

NOMENCLATURE

k_1	Kinetics reaction constants	hr^{-1}
k_2 and k_3	Kinetics reaction constants	$m^3/(mol\,hr)$
P	Pressure of the reactor	(Pa)
R	Constant of gases	J/(mol K)
t	Time	(hr)
t_f	Final time	(hr)
t_s	Switching time	(hr)
T	Temperature	(K)
u	Pure feed rate of A	(mol/hr)
V	Volume of the reactor	(m^3)
x	Vector of state variables	(-)

Greek Letters

λ	Vector of costate variables	(-)
δ	Vector of slack variables	(-)

Subscripts

max	Maximum	(-)
min	Minimum	(-)
0	Initial	(-)

REFERENCES

Ascher, U., Spiteri, R., Collocation Software for Boundary-Value Differential-Algebraic Equations. SIAM J. Scient. Comput., 15, 938-952 (1994).

Bellman, R. E., Dynamic Programming. Princeton University Press (1957).

Biegler, L. T., Cervantes, A. M., Wachter, A., Advances in Simultaneous Strategies for Dynamic Process Optimization. Chemical Engineering Science, 57, 575-593 (2002).

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

Bulirsch, R., Montrone, F., Pesch, H. J., Abort Landing in the Presence of a Windshear as a Minimax Optimal Control Problem, Part I: Necessary Conditions. Journal of Optimization Theory and Applications, 70, 1-23 (1991a).

Bulirsch, R., Montrone, F., Pesch, H. J., Abort Landing in the Presence of a Windshear as a Minimax Optimal Control Problem, Part II: Multiple Shooting and Homotopy. Journal of Optimization Theory and Applications, 70, 223-254 (1991b).

Bulirsch, R., Nerz, E., Pesch, H. J., Combining Direct and Indirect Methods in Optimal Control: Range Maximization of a Hang Glider. International Series of Numerical Mathematics, 111, 273-288 (1993).

Cervantes, A., Biegler, L. T., Optimization Strategies for Dynamic Systems. In C. Floudas, P. Pardalos (Eds), Encyclopedia of Optimization, Dordrecht: Kluwer Academic Publishers (1999).

Feehery, W. F., Dynamic Optimization with Path Constraints. Ph.D. diss, Massachusetts Institute of Technology, USA (1998).

Gerdts, M., A Direct Shooting Method for the Numerical Solution of Higher Index DAE Optimal Control Problems. Copyright by the author (2001).

Gomes, E. O., Abordagem Algébrico-Diferencial na Solução de Problemas de Controle Ótimo. M.Sc. thesis, PEQ/UFU, Federal University of Uberlândia (2000).

Huang, Y. J., Reklaitis, G. V., Venkatasubramanian, V., Model Decomposition Based Method for Solving General Dynamic Optimization Problems. Computers and Chemical Engineering, 26, 863-873 (2002).

Jacobson, D. H., Lele, M., A Transformation Technique for Optimal Control Problems with a State Variable Inequality Constraint. IEEE Transactions on Automatic Control, 14, 457-464 (1969).

Koslik, B., Breitner, M. H., In Optimal Control Problem in Economics with Four Linear Controls. Journal of Optimization Theory and Applications, 94, 619-634 (1997).

Lobato, F.S., Abordagem Mista para Problemas de Otimização Dinâmica. M.Sc. thesis, PEQ/UFU, Federal University of Uberlândia (2004).

Pontryagin, L. S., Boltyanskil, V. G., Gamkrelidge, R. V. Mishchenko, E. F., The Mathematical Theory of Optimal Process. New York: Interscience (1962).

Ray, W. H., Advanced Process Control. New York: McGraw-Hill (1981).

Santos, K. G., Lobato, F. S., Ribeiro, E. J., Oliveira-Lopes, L. C., Murata, V. V., Controle Ótimo da Fermentação Alcoólica com Altas Concentrações Iniciais de Substrato em Reator Batelada Alimentada. In CD Rom of XV SINAFERM, Florianópolis – Brasil (2005).

Stryk, O. Von., User's Guide for DIRCOL – A Direct Collocation Method for the Numerical Solution of Optimal Control Problems. Technische Universitat Darmastadt, Fachgebiet Simulation and Systemoptimierung (SIM) (1999).

Stryk, O. Von, Bulirsch, R., Direct and Indirect Methods for Trajectory Optimization. Annals of Operations Research, 37, 357-373 (1992).

Unger, J., Kroner, A., Marquardt, W., Structural Analysis of Differential Algebraic Equations Systems: Theory and Applications. Computers and Chemical Engineering, 19, 867-882 (1995).