Optimal Control of a Chemical Reactor

Jan Cvejn

Abstract
In the paper we describe a mathematical model of a chemical reactor with concentrated parameters and show some results obtained by dynamic optimization of the inlet and the outlet flows of entering components and the mixture. The flows are considered as functions of time and are subject to real constraints. Suitable criterion of optimality can be constructed from the history of concentration of the product and the history of the control variables. The results obtained by numerical computation show that for some choices of the criterion meaningful optimal trajectories in the state space exist. Control along such trajectories can increase quality of the product by means of control logic, with low acquisition and operational costs.

Keywords: Optimal control, optimization, chemical processes

Introduction

Chemical technology uses two basic types of reactors: batch and flow ones. An advantage of batch-type reactors, which realize a reaction of finite volume of inlet components in a sufficiently large time horizon, is high degree of conversion of inlet components. A disadvantage for larger volumes of production per time unit is big capacity of the reactor required, which significantly influences both acquisition and operational costs and cannot be increased arbitrarily. Larger volumes of production under limited cost of equipment can be achieved using flow reactors with smaller capacity, which however work with lower degree of conversion.

Besides a change of device parameters and consequent increasing acquisition costs there however exist other possibilities of enhancing the parameters of production. Using suitable changes of control inputs in time, over a finite time horizon, increasing the concentration of the product can be achieved as well. This holds especially in the case of parallel or bi-directional reactions, where desired product of the reaction can further react in an undesirable way.

Consider a chemical reaction of type

$$\alpha_1 A_1 + ... + \alpha_m A_m \rightarrow \beta_1 B_1 + ... + \beta_n B_n \,.$$

If c_{Ai}, c_{Bj} $[\mathrm{mol}.m^{-3}]$ are the concentrations of reacting components and products, c_{0i} $[\mathrm{mol}.m^{-3}]$ the concentrations of inlet components A_i, V $[m^3]$ volume of the mixture in the reactor, q_i $[m^3.s^{-1}]$ the inlet flows and q $[m^3.s^{-1}]$ the outlet flow, behavior of the reaction is described by a set of nonlinear equations in the form [4], [6]:

$$\begin{aligned} \frac{d(Vc_{Ai})}{dt} &= -\alpha_i r_T V \prod_{k=1}^{m} c_{Ak}^{\alpha_k} + q_i c_{0i} - q c_{Ai}, \\ \frac{d(Vc_{Bj})}{dt} &= \beta_j r_T V \prod_{k=1}^{m} c_{Ak}^{\alpha_k} - q c_{Bj}, \end{aligned} \tag{1}$$

$$\frac{dV}{dt} = \sum_{i=1}^{m} q_i - q \tag{2}$$

for all $i = 1,...,m$, $j = 1,...,n$. We assume for simplicity that the reaction is isothermic, so the reaction rate term r_T is constant. Otherwise, it depends on the temperature according to the formula

$$r_T = r_0 e^{\left(\frac{-E}{RT}\right)} \tag{3}$$

where r_0, E, R are constants.

If we put

$$\begin{aligned} \mathbf{u} &= (q_1,...,q_m,q)^T \\ \mathbf{x} &= (c_{A1},...,c_{Am},c_{B1},...,c_{Bn},V)^T \end{aligned} \tag{4}$$

we obtain a nonlinear system in the state-space description form

$$\dot{\mathbf{x}}(t) = \mathbf{f}\left(\mathbf{x}(t),\mathbf{u}(t)\right). \tag{5}$$

Values of the control variables are subject to real constraints

$$u_i \in \left[0, u_{i\max}\right]. \tag{6}$$

All the state variables x_i can be subject to constraints given by technological requirements as well.

A general form of the criterion to be minimized is [2]:

$$J = \varphi(\mathbf{x}(t_f)) + \int_0^{t_f} f_0\left(\mathbf{x}(t),\mathbf{u}(t),t\right) dt \rightarrow \min \tag{7}$$

where continuously differentiable functions φ and f_0 are chosen to express practical demands on the process, t_f is the time of running the reaction. The criterion (7) is to be minimized over the space of functions $\mathbf{u}(t)$ with respect to

dynamic and algebraic constraints (5) and (6). The problem can be complicated by ending conditions in the form

$$\boldsymbol{\psi}\left(\mathbf{x}(t_f)\right) = 0\,. \tag{8}$$

where $\boldsymbol{\psi}$ is a continuously differentiable vector function.

Necessary conditions of optimality for the problem (5) - (8) are known as Pontryagin's principle of maximum [1], [8]. If $H(\mathbf{x},\mathbf{u},\boldsymbol{\lambda},t) = f_0 + \boldsymbol{\lambda}^T\mathbf{f}$ is the Hamiltonian with the state adjoint function $\boldsymbol{\lambda}(t)$, for optimal trajectory $\mathbf{x}(t)$ holds:

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x},\mathbf{u}^*) \tag{9}$$

$$\mathbf{x}(0) = \mathbf{x}_0\,,\quad \boldsymbol{\psi}\left(\mathbf{x}(t_f)\right) = 0 \tag{10}$$

$$\dot{\boldsymbol{\lambda}}^T = -\frac{\partial H(\mathbf{x},\mathbf{u}^*,\boldsymbol{\lambda},t)}{\partial \mathbf{x}} \tag{11}$$

$$\boldsymbol{\lambda}^T(t_f) = \left(\frac{\partial \varphi}{\partial \mathbf{x}} + \boldsymbol{\nu}^T\frac{\partial \boldsymbol{\psi}}{\partial \mathbf{x}}\right)\Bigg|_{t=t_f} \tag{12}$$

$$\mathbf{u}^*(t) = \arg\min_{\mathbf{u}\in U}\left\{H(\mathbf{x},\mathbf{u},\boldsymbol{\lambda},t)\right\} \tag{13}$$

where U is the set of feasible values of $\mathbf{u}$ (6) and $\boldsymbol{\nu}$ is a Lagrange multipliers vector.

1. The Criterion of Optimality

There are several reasonable choices of f_0 and φ that come into account. In all the cases we consider that the time of reaction t_f is given. A more complicated situation can be dealt with, where t_f is not known and is to be optimized.

One possibility is to maximize the volume or the concentration of a product B_k, subject to given amount of inlet components and the constraints (6):

$$J = \int_0^{t_f} -q(t)\,c_{Bk}(t)\,dt \to \min \tag{14}$$

$$\int_0^{t_f} u_i(t)\,dt = W_i\,. \tag{15}$$

Instead of the constraints (15) on the volumes of inlet components it is possible to consider a single constraint on total volume of the mixture produced:

$$\int_0^{t_f} q(t)\,dt = W\,. \tag{16}$$

The isoperimetric constraints (15) or (16) can be replaced by trivial ending conditions $y_i(t_f) = W_i$ or $y(t_f) = W$, respectively, by extending the state-space description of equations $\dot{y}_i = u_i$ or $\dot{y} = q$. Initial conditions are trivial in most cases: $x_i(0) = x_{i0}$.

If we substitute (14) into the necessary conditions of optimality (9) - (13), the Hamiltonian H is affine in $\mathbf{u}$. It follows that if the Hamiltonian is dependent on all the components of $\mathbf{u}$, the extremal $\mathbf{u}^*(t)$ must be for each t on the boundary of U. Either $\mathbf{u}^*(t)$ is constant or its components are discontinuous at some time points in $\left[0,t_f\right]$ and change from 0 to $u_{i\max}$ or back. It can however happen that in a certain interval $H(\mathbf{x},\mathbf{u},\boldsymbol{\lambda},t)$ is independent of some components of $\mathbf{u}$. In such a case the problem can have so-called singular solution [2].

Although the optimal control history will probably contain step changes, discontinuities in the control flows can be disadvantageous in practice. Continuous histories that are an approximation of the actual discontinuous solution are obtained if the criterion is extended of quadratic terms of control:

$$J = \int_0^{t_f}\left[-q(t)\,c_{Bk}(t) + \beta\sum_i u_i^2\right]dt\,. \tag{17}$$

It seems that another advantage of quadratic terms in the criterion is in many cases faster convergence to the solution.

2. The Optimization Method

There exist numerical methods based directly on the Pontryagin's maximum principle [7], [9]. In this paper an alternative approach is described. For relatively simple, but approximate numerical solution the algebraic constraints (6) and the terminal conditions (8) can be removed and added to the criterion by means of smooth penalty terms [2], [7]. The transformed functional has the form

$$J' = \varphi(\mathbf{x}(t_f)) + \varphi_p(\mathbf{x}(t_f)) + \\ + \int_0^{t_f}\left[f_0(\mathbf{x},\mathbf{u},t) + f_p(\mathbf{u})\right]dt \to \min \tag{18}$$

where

$$f_p(\mathbf{u}) = \sigma\sum_{j=1}^{\dim u}\left[\min\{u_i,0\}^2 + \max\{u_i - u_{i\max},0\}^2\right] \tag{19}$$

$$\varphi_p(\mathbf{x}(t_f)) = \sigma\left\|\boldsymbol{\psi}(\mathbf{x}(t_f))\right\|^2 \tag{20}$$

and $\sigma > 0$ is sufficiently large, but not too large. Similar terms can be used to replace the state constraints.

To obtain a solution with higher precision, a sequence of solutions for rising σ has to be constructed, where the previous solution is taken as initial estimate for the next step, but it is well known that this method is rather inefficient [5].

To solve the problem with augmented criterion (18) we used an approach described in more detail in [3]. Let us choose a value of parameter M and define the set of continuous functions such that the components of the optimal control function $\mathbf{u}(t)$ can be in interval $\left[0,t_f\right]$ well approximated by weighted sums

$$u_j(t) = \sum_{k=1}^{M} a_{jk}\,b_k(t) \tag{21}$$

where $j = 1,\ldots,m$ and the base functions $\{b_k(t)\}$, $i = 1,\ldots,M$ are independent. Expression (21) can be written in matrix notation as

$$\mathbf{u}(t) = \mathbf{A}\mathbf{b}(t) \tag{22}$$

where $\mathbf{A}$ is a matrix of weighting coefficients. If we substitute (22) into (5) and (7), we obtain a set of coupled differential equations

$$\frac{d}{dt}\begin{bmatrix} J \\ \mathbf{x}\end{bmatrix} = \begin{bmatrix} f_0(\mathbf{x}(t),\mathbf{A}\mathbf{b}(t),t) \\ \mathbf{f}(\mathbf{x}(t),\mathbf{A}\mathbf{b}(t),t)\end{bmatrix} \tag{23}$$

with initial conditions

$$\begin{bmatrix} J(0) \\ \mathbf{x}(0) \end{bmatrix} = \begin{bmatrix} 0 \\ \mathbf{x}_0 \end{bmatrix}. \tag{24}$$

Equation (23) can be numerically integrated from 0 to t_f for given $\mathbf{A}$ to obtain the criterion value $J(t_f) = J(\mathbf{A})$. The parameter space is thus discretized. The value $J(\mathbf{A})$ then can be numerically minimized to obtain the optimal coefficients $\mathbf{A}^* = (a_{jk}^*)$.

Consider that the components of $\mathbf{A}$ are reordered into a vector $\mathbf{z}$, where $\dim \mathbf{z} = m.M$. For minimization of $J(\mathbf{A}) = J(\mathbf{z})$ we are using the quasi-Newton algorithm

$$\mathbf{z}_{k+1} = \mathbf{z}_k - \alpha \mathbf{G}_k \nabla_k \tag{25}$$

where ∇_k is the gradient of $J(\mathbf{z})$ in $\mathbf{z}_k$, $\alpha \geq 0$ is the value obtained by line-search minimization of $J(\mathbf{z})$ from $\mathbf{z}_k$ in direction $-\mathbf{G}_k \nabla_k$. Matrix $\mathbf{G}_k$, which represents the approximation of the inverse of Hessian $\left(\partial^2 J / \partial z_i \partial z_j\right)$ in $\mathbf{z}_k$, is set as unit matrix in the first step and is updated in each step using the BFGS formula [5]:

$$\mathbf{G}_{k+1} = \mathbf{G}_k + \left(1 + \frac{\boldsymbol{\gamma}^T \mathbf{G}_k \boldsymbol{\gamma}}{\boldsymbol{\delta}^T \boldsymbol{\gamma}}\right) \frac{\boldsymbol{\delta}\boldsymbol{\delta}^T}{\boldsymbol{\delta}^T \boldsymbol{\gamma}} - \frac{\boldsymbol{\delta}\boldsymbol{\gamma}^T \mathbf{G}_k + \mathbf{G}_k \boldsymbol{\gamma}\boldsymbol{\delta}^T}{\boldsymbol{\delta}^T \boldsymbol{\gamma}} \tag{26}$$

where $\boldsymbol{\delta} = \mathbf{z}_{k+1} - \mathbf{z}_k$, $\boldsymbol{\gamma} = \nabla_{k+1} - \nabla_k$.

The gradient $\nabla J(\mathbf{z}) = \left(\partial J / \partial z_1, ..., \partial J / \partial z_{Mm}\right)^T$ of the criterion can be in the beginning approximated by central differences

$$\frac{\partial J}{\partial z_i} \approx \frac{J(\mathbf{z} + h\mathbf{e}_i) - J(\mathbf{z} - h\mathbf{e}_i)}{2h} \tag{27}$$

where $h > 0$ is a small number and $\mathbf{e}_i$ the unit vector in i-th direction. Since every computation of $J(\mathbf{A})$ needs integration of (23), the evaluation of (27) can be a considerably time-consuming operation for larger M. In [3] it is shown that exact formula for the gradient of $J(\mathbf{A})$ is

$$\frac{\partial J}{\partial a_{ij}} = \int_0^{t_f} \frac{\partial H}{\partial u_i} b_j(t)\, dt \tag{28}$$

where $H(\mathbf{x}, \mathbf{u}, \boldsymbol{\lambda}, t)$ is the Hamiltonian. Computation using this formula requires explicit knowledge of partial derivatives $(\partial H / \partial \mathbf{x})$ and $(\partial H / \partial \mathbf{u})$, but in the case of models of type (1) this information is usually easily available. Demand factor of determination of ∇J can be significantly reduced for special types of base systems as shown in [3]. In the example below we chose a sufficiently large N and used the simplest - piecewise constant base system defined for $i = 1, ..., N$ as

$$b_i(t) = \eta(t - (i-1)T) - \eta(t - iT), \tag{29}$$

where $\eta(t)$ is Dirac unit-step function.

3. Solved Example

For a demonstration consider the case of a reaction of type $2A + B \rightarrow C$ (Fig. 1). The reaction is isothermic and the reactor contents is perfectly mixed.

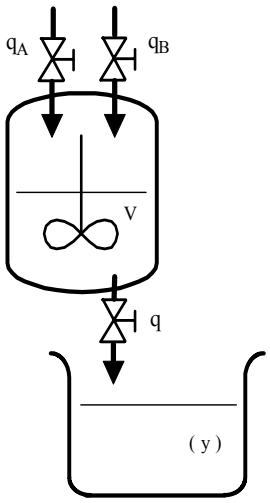

Fig.1 A reactor example

By substitution into equations (1) - (2) we obtain

$$\begin{aligned} \frac{d(Vc_A)}{dt} &= q_A c_{A0} - 2r_T V c_A^2 c_B - q c_A \\ \frac{d(Vc_B)}{dt} &= q_B c_{B0} - r_T V c_A^2 c_B - q c_B \\ \frac{d(Vc_C)}{dt} &= r_T V c_A^2 c_B - q c_C \\ \frac{dV}{dt} &= q_A + q_B - q. \end{aligned} \tag{30}$$

For $k = A, B, C$ holds

$$\begin{aligned} \frac{d(Vc_k)}{dt} &= V \frac{dc_k}{dt} + c_k \frac{dV}{dt} = \\ &= V \frac{dc_k}{dt} + c_k (q_A + q_B - q). \end{aligned} \tag{31}$$

The equations (30) then can be rewritten into the standard form

$$\begin{aligned} \frac{dc_A}{dt} &= \frac{1}{V}\left[q_A c_{A0} - (q_A + q_B) c_A\right] - 2r_T c_A^2 c_B \\ \frac{dc_B}{dt} &= \frac{1}{V}\left[q_B c_{B0} - (q_A + q_B) c_B\right] - r_T c_A^2 c_B \\ \frac{dc_C}{dt} &= \frac{1}{V}\left[-(q_A + q_B) c_C\right] + r_T c_A^2 c_B \\ \frac{dV}{dt} &= q_A + q_B - q. \end{aligned} \tag{32}$$

The control and the state vector are

$$\begin{aligned} \mathbf{u} &= (q_A, q_B, q)^T \\ \mathbf{x} &= (c_A, c_B, c_C, V)^T. \end{aligned} \tag{33}$$

The parameter values were chosen as:

$$\begin{aligned} q_{A\max} &= q_{B\max} = 10^{-3}\, m^3.s^{-1} \\ q_{\max} &= 1.5 \times 10^{-3}\, m^3.s^{-1} \\ V_{\max} &= 0.1 m^3 \\ c_{A0} &= c_{B0} = 1 \text{mol}.m^{-3} \\ r_T &= 1 \\ t_f &= 300 s. \end{aligned}$$

The criterion to be minimized is

$$J = \int_0^{t_f} \left(-10\,q.c_C + q_A^2 + q_B^2 + q^2\right) dt \qquad (34)$$

with the isoperimetric constraint on total volume

$$\int_0^{t_f} q(t)\,dt = W = 0.3\,m^3\,. \qquad (35)$$

The constraint (35) was eliminated by extending the state description of the equation $\dot{y} = q(t)$.

The initial conditions are

$$c_A(0) = c_B(0) = c_C(0) = 10^{-3}\ \mathrm{mol}.m^{-3}$$
$$V(0) = 0.01\ m^3,\ y(0) = 0\,m^3\,.$$

The terminal conditions are

$$V(t_f) = 0.01\ m^3$$
$$y(t_f) = W = 0.3\,m^3\,.$$

The figures 2 - 4 below show obtained histories of the state and the control variables. The computation is rather time-demanding. For approximation of the control history $N = 32$ was chosen. Total number of optimized parameters is 99.

The example shows that meaningful optimal trajectory exists. The optimal process obviously consists of three phases. In the middle phase, where all three control valves are opened, the flows are constant. In the initial and the ending phase some control variables reach their maximal or minimal values.

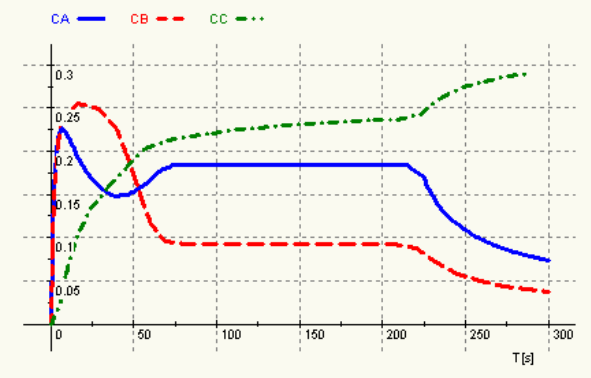

Fig.2 The histories of the concentrations
$c_A, c_B, c_C\ [\mathrm{mol}.m^{-3}]$

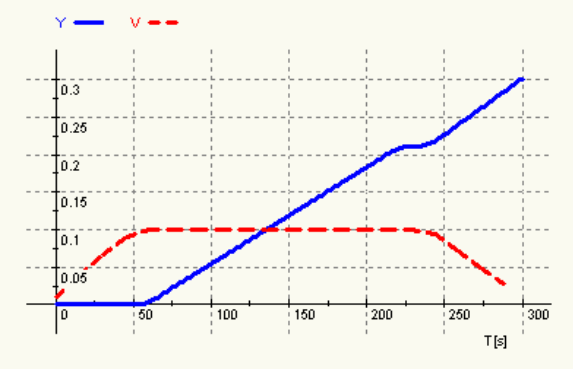

Fig.3 The histories of the volumes $V(t), y(t)[m^3]$

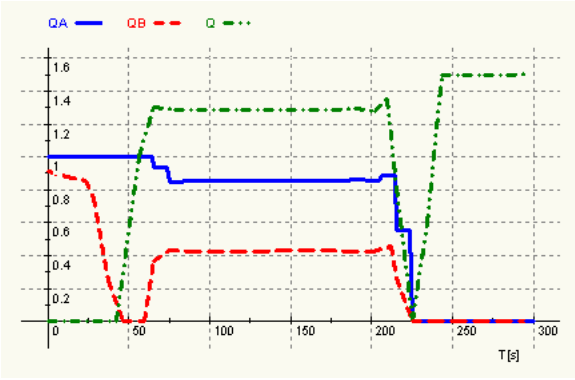

Fig.4 The histories of the control variables
$q_A, q_B, q\ [10^{-3} m^3.s^{-1}]$

Using the knowledge of optimal history the quality of the product can be enhanced by means of control logic. Another way how to increase concentration of the product is using a rector with bigger capacity or connecting more reactors into series, which can be however much more expensive a solution.

References

[1] ALEXEJEV, V. M. et al.: Matematická teorie optimálních procesů. Academia. Praha, 1991.

[2] BRYSON, A. E., HO, Y. C.: Applied Optimal Control. Hemisphere Publishing Corp. New York, 1975.

[3] CVEJN, J.: Dynamically Optimal Trajectories Design - Generalized Discretization Method. Proc. 13th IEEE IFAC International Conference on Methods and Models in Automation and Robotics. Szczecin, Poland, 2007. (*In print*)

[4] ČERMÁK, J., PETERKA, V., ZÁVORKA, J.: Dynamika regulovaných soustav v tepelné energetice a chemii. Academia. Praha, 1968.

[5] FLETCHER, R.: Practical Methods of Optimization. Second Edition. John Wiley & Sons Ltd., 2001.

[6] KOMŮRKA, J., GEMZA, E.: Matematické modelování procesů chemické technologie. Vysoká škola chemicko-technologická. Pardubice, 1976.

[7] POLAK, E.: Computational Methods In Optimization. Academic Press, New York, 1971.

[8] PONTRYAGIN, L. S. et al.: The Mathematical Theory of Optimal Processes. Interscience Publishers Inc. New York, 1962.

[9] VASILJEV, F.: Čislennyje metody rešenija extremalnych zadač. Nauka. Moskva, 1988.

Doc. Ing. Jan Cvejn, Ph.D.

University of Pardubice
Institute of Electrical Engineering and Informatics
Studentská 95
532 10 Pardubice, Czech Republic
E-mail: jan.cvejn@upce.cz