

HOW TO AVOID OVERCOOKING: OPTIMAL YIELD OF B FROM THE CONSECUTIVE CHEMICAL REACTIONS $A \rightarrow B \rightarrow C$

Thor A. Bak¹, Bjarne Andresen^{2,3}, James D. Nulton², Benjamin Felts III², and Peter Salamon²

¹Chemistry Laboratory III, University of Copenhagen, Universitetsparken 5,
DK 2100 Copenhagen Ø, Denmark

²Department of Mathematical Sciences, San Diego State University,
San Diego, CA 92182 7720, USA

³Ørsted Laboratory, University of Copenhagen, Universitetsparken 5,
DK 2100 Copenhagen Ø, Denmark

ABSTRACT

We analyze the consecutive chemical reactions $A \rightarrow B \rightarrow C$ both numerically and analytically using temperature as the control variable. Starting with pure A and maximizing the yield of B at the end of the given process duration, all optimal paths start with a branch at infinite temperature. We derive the switching curve along which jumps are allowed from this temperature to lower temperatures. We also find that the marginal productivity of time equals the final production rate, i.e., the incremental gain of B when the path is re-optimized for increased duration equals the production rate of B at the end of the reaction. We further find that for given parameters there is a unique "maximum useful time" which results in the largest yield of B possible. For this duration to be finite, it is necessary that the smallest activation energy be one of the activation energies for B reacting. If a duration longer than this is specified, all reactions should be shut off for that excess amount of time, a situation which makes most optimization routines become unstable.

Keywords: Optimal control, Chemical reactions, Temperature control, Optimal yield

INTRODUCTION

Already half a century ago, several authors [1-11] considered the question of improving the yield of a chemical reaction taking place in a tubular reactor by adjusting the temperature in various parts of the reactor. The papers most pertinent to the present work are those of Bilous and Amundson [3] and Aris [8-11]. Since then, numerical and approximate techniques have become widespread in the literature. Analytical approaches to understanding the structure of possible solutions have met with grave difficulties on the other hand and have been abandoned in favor of a useful engineering approach which can solve specific real-world problems of industrial interest. This circumstance in which the search for general principles is abandoned in favor of the explicit solution of practical problems is reminiscent of the state of energy conversion technology in the early seventies when R.S. Berry

and others [12-13] approached thermodynamics with the aim of extracting general principles for what is possible given the constraint of finite time. There too, practical engineering optimizations were technically up to the task of designing real applications but there was a lack of general results concerning limits to what is possible in finite-time. This paper is offered in a similar spirit. We strive to increase the analytic understanding of the possible control of chemical reactions for maximal yield with a focus on finite time.

By considering a reaction in a tubular reactor through which the reacting material is pumped and along which the temperature varies, we are in effect considering a reaction at a temperature which varies with time. If there is only one chemical reaction going on, the more or less obvious answer to the problem of getting maximum yield is to adjust the temperature so that the reaction rate is as large as possible at each

point of the reactor. In the following, we let the temperature be infinite, but this just means "as high as possible, other conditions taken into account".

If there are two competing reactions $A \xrightarrow{k_1} B$ and $A \xrightarrow{k_2} C$ for which B is the desired product and C is a waste product, the situation is not so simple, since by increasing the rate constants for the first reaction we automatically also increase the rate constants of the second thereby letting A go directly to waste. If there are two consecutive reactions $A \xrightarrow{k_3} B \xrightarrow{k_4} C$ for which as before B is the desired product and C the waste, the situation is additionally complicated and furthermore so if the reverse reactions are also taken into account.

The latter system is a crude model for the operation of a tunnel oven where B is the desired product and C is deadburned material. Thus for gypsum A is $\text{CaSO}_4 \cdot 2 \text{H}_2\text{O}$, B is burned gypsum $\text{CaSO}_4 \cdot 1/2 \text{H}_2\text{O}$ which readily reabsorbs water under solidification, while C is deadburned gypsum CaSO_4 which only very slowly reacts with water. Another example is the production of cement. The actual reactions going on in an industrial oven are quite complicated and we will not attempt to reproduce them. Our model only describes the general features using relative values of rate constants and activation energies to exemplify the phenomenon.

Aris [8] solved the problem of two consecutive reactions using the method of dynamic programming rather than functional differentiation used by Bilous and Amundson [3]. He thereby also resolved two questions left open by them. In this paper we use the method of optimal control theory, on the problem $A \xrightarrow{k_1} B \xrightarrow{k_2} C$. Our analysis confirms the result of Amundson and Aris that one should not merely adjust the temperature to its highest value throughout the reaction. At a certain stage, when the back reaction begins to be important, the temperature must be lowered. We analyze this feature in detail. The method we use is more powerful than those used previously and it may be applied to other and more complicated reaction schemes. However, except when further simplifications are introduced (as shown in Section 5), an analytical solution cannot be obtained.

SYSTEM

We consider the system of first order chemical reactions



with rate constants k_i , $i=1\dots 4$ depending on temperature T through Arrhenius expressions,

$$k_i = s_i e^{-E_i / kT} \quad (2)$$

with activation energies E_i and collision factors s_i . These are taken to be constants whereas the temperature T varies with time because, as mentioned in the introduction, the reacting mixture is pumped through a tube along which the temperature varies. From now on we forget the experimental setup and just assume we have a reaction mixture for which the temperature can vary with time and take this to be our control variable.

Starting from pure A, i.e. $[A]=1$, $[B]=[C]=0$, our objective is to produce as much B as possible within a given duration Δ by varying the temperature appropriately over the time span $[0, \Delta]$. We will not put any restrictions on the temperature except that it must be positive, and we assume that it can be changed freely and instantaneously as needed.

THE FORMAL PROBLEM

We use optimal control theory [14] to derive the optimal temperature path. In order to avoid the omnipresent exponential functions in the rate constants, it will prove convenient to change the control variable from T to

$$u = e^{-1/kT} \quad (3)$$

with u restricted to the interval $[0,1]$ corresponding to the temperature interval $[0, \infty]$ so that

$$k_i = s_i u^{E_i} \quad (4)$$

Our objective function to be maximized is thus

$$B(\Delta) = \int_0^\Delta \frac{dB}{dt} dt \quad (5)$$

subject to the three constraints

$$\frac{dA}{dt} = \lambda_1 A + k_2 B \quad (6)$$

$$\frac{dB}{dt} = k_1 A - k_2 B - k_3 B + k_4 C \stackrel{def}{=} F \quad (7)$$

$$A + B + C = 1 \quad (8)$$

and with initial conditions

$$A(0) = 1, \quad B(0) = C(0) = 0. \quad (9)$$

Here and in the following we have skipped the conventional square brackets when writing concentrations, i.e. we use e.g. A instead of $[A]$. The third rate equation dC/dt is a consequence of dA/dt , dB/dt and mass conservation, eq. (8) and thus automatically satisfied.

We will at once reduce the set of equations further by substituting C from eq. (8) into eq. (7) so that the optimization only has two constraints, eq. (6) and

$$\frac{dB}{dt} = (k_1 - k_4)A - (k_2 + k_3 + k_4)B + k_4. \quad (10)$$

The Hamiltonian for this system is

$$H = \lambda_0 [(k_1 - k_4)A - (k_2 + k_3 + k_4)B + k_4] + \lambda_1 (\lambda_1 A + k_2 B) + \lambda_2 [(k_1 - k_4)A - (k_2 + k_3 + k_4)B + k_4] \quad (11)$$

where the conjugate variable λ_0 associated with the objective function must be constant and can only have the values 0 or 1. However, since the final values of A and B are free so that $\lambda_0 = \lambda_1 = 0$, $\lambda_0 = 0$ is not a possibility as it would make the final value of the total conjugate vector equal to zero. Upon combining the two Lagrange multipliers λ_0 and λ_1 into a single $\lambda = \lambda_0 + \lambda_1$, the Hamiltonian becomes

$$H = \lambda (\lambda_1 A + k_2 B) + \lambda [(k_1 - k_4)A - (k_2 + k_3 + k_4)B + k_4] \quad (12)$$

The Hamiltonian equations to be solved are then

$$\frac{\partial H}{\partial A} = \lambda \frac{d\lambda}{dt} \quad \frac{\partial H}{\partial B} = \lambda \frac{d\lambda}{dt} \quad (13)$$

$$\frac{\partial H}{\partial \lambda} = \frac{dA}{dt} \quad \frac{\partial H}{\partial \lambda} = \frac{dB}{dt}$$

under the boundary conditions eq. (9) and

$$\lambda(0) = 0 \quad \lambda(1) = 1. \quad (14)$$

Thus we have initial conditions on the state variables but final conditions on the conjugate variables, corresponding to the final states being free to vary to achieve the optimality. The two lower equations in (13) which specify the time evolution of the state variables A and B are already stated above as eqs. (6) and (10). Finally, the control $u(t)$ is determined so as to maximize H :

$$u = \arg \max_u H, \quad (15)$$

where $\arg \max H$ means the value of the argument u of H which results in the maximum possible value of H . The Hamiltonian H depends on the "temperature" u through the rate constants k_i .

SOME CONSEQUENCES OF OPTIMALITY

Our first observation is valid much more generally than the example considered here. We first note that the Hamiltonian in any problem without explicit time dependence except for a constrained total time λ equals the incremental yield of the objective function by increasing the duration of the process,

$$H = \frac{dB^*}{d\lambda}, \quad (16)$$

where B^* is the optimal value of the objective function B over possible controls lasting total time λ . It then follows that when we are maximizing the yield of one chemical, say B , (the value of one state variable) without regard for the final values of other species (unconstrained final state variables), the marginal yield per unit of time spent is exactly equal to the rate of B production at that instant, i.e. $dB^*/d\lambda = dB/dt$. This follows by substituting the boundary conditions eq. (14) into the Hamiltonian expression eq. (12) and identifying the resulting expression as dB/dt ,

eq. (10). This observation is true generally, provided the objective function coincides with the final value of one of the state variables, and none of the other state variables have constrained final values, a common occurrence for chemical systems.

Although our ultimate aim is primarily to understand the effects of finite time limitations, consider for a moment the problem with unconstrained duration τ . For this case, the Hamiltonian must vanish and so will the final reaction rate. Suppose there is a finite optimal duration τ^* . Then any duration $\tau > \tau^*$ is also optimal since we can assure the same yield by proceeding as before up to time τ^* and then setting the temperature to 0 ($u=0$). Thus the structure of such problems is highly unusual in the sense that they have a *maximal useful time* τ_{mu} .

Any time allocated beyond τ_{mu} is spent at 0 temperature with the reactions shut off. Note that $u=0$ implies that the Hamiltonian is zero and all of the dynamical equations give zero derivatives – as expected on physical grounds, the process is frozen. Such $u=0$ branches can be added to any optimal process and still have the results satisfy all the conditions of optimality. In fact for solutions corresponding to $H=0$, such branches can be spliced in anywhere along the process once the constrained time τ exceeds the maximum useful time, $\tau_{\text{mu}} = \min\{\tau^*\}$. This fact serves to confound numerical algorithms for the solution as described further below.

The finite τ_{mu} case occurs only when the desired product is unstable in the sense that the activation energy for B reacting to form A or C is the lowest of the four activation energies in the system, i.e. $\min\{E_1, E_2, E_3, E_4\} = E_2$ or E_3 . If this is not the case, we can always squeeze out more yield by lowering the temperature. To see this, consider the situation at a particular set of concentrations A, B, and C. The rate of B formation is $k_1 A + k_4 C$ while the rate of B removal is $(k_2 + k_3)B$. Suppose E_1 is the smallest. To facilitate comparison between B production and B removal we divide both by k_1 yielding $A + k_4/k_1 C$ and

$\frac{k_2 + k_3}{k_1} B$. All ratios in these expressions have a

factor of $u^{E_i/\tau E_1}$ and since E_1 is the smallest, these can be made arbitrarily small by lowering the temperature. It thus follows that if $\tau_{\text{mu}} < \tau$, E_1 cannot be the lowest activation energy. A similar argument holds for the case when E_4 is the lowest. When τ_{mu} is infinite, the maximum yield B^* is approached asymptotically. In fact, when E_1 and E_4 are the two lowest activation energies in the system, we can convert to 100% product in infinite time by progressively lowering the temperature as argued above.

NUMERICAL OPTIMIZATION

The $H=0$ solutions are unstable and this implies that small H solutions are also unstable unless we work from a nearby point. In practice, this can be achieved by solving progressively longer version of the problem and starting from the previous solution.

For the numerical examples of this boundary value problem we solve eqs. (13) using Krotov's method [15] in which alternately the state differential equations (in A and B) are solved in the forward time direction and the conjugate differential equations (in λ and μ) are solved in the backward direction. In this way each set of equations starts with its given boundary conditions. After each iteration back and forth the u which maximizes H is determined at each point along the trajectory. This method is much more efficient than the conventional "shooting method" and usually converges in just a few iterations.

The optimization converges quickly as long as τ is not too long, i.e. as long as time is at a premium so that it must be used carefully. For long reaction times B will react further to become C, depleting the desired product. This can be prevented only by shutting off all reactions by reducing the temperature to zero. However, these periods of standstill can equally well be placed anywhere during the process, either as one large period or as several shorter ones, leaving the optimization in a dilemma as to what to do. Mathematically this shows up as unstable equations which cannot converge but keep switching between zero and

non-zero temperature. In these situations the Hamiltonian has multiple minima for u in the time interval, and small variations during the iterative solution may easily flip the control between very

different values, causing erratic behavior. We were only able to obtain numerical solutions for duration $\Delta t < \Delta t_{\min}^*$.

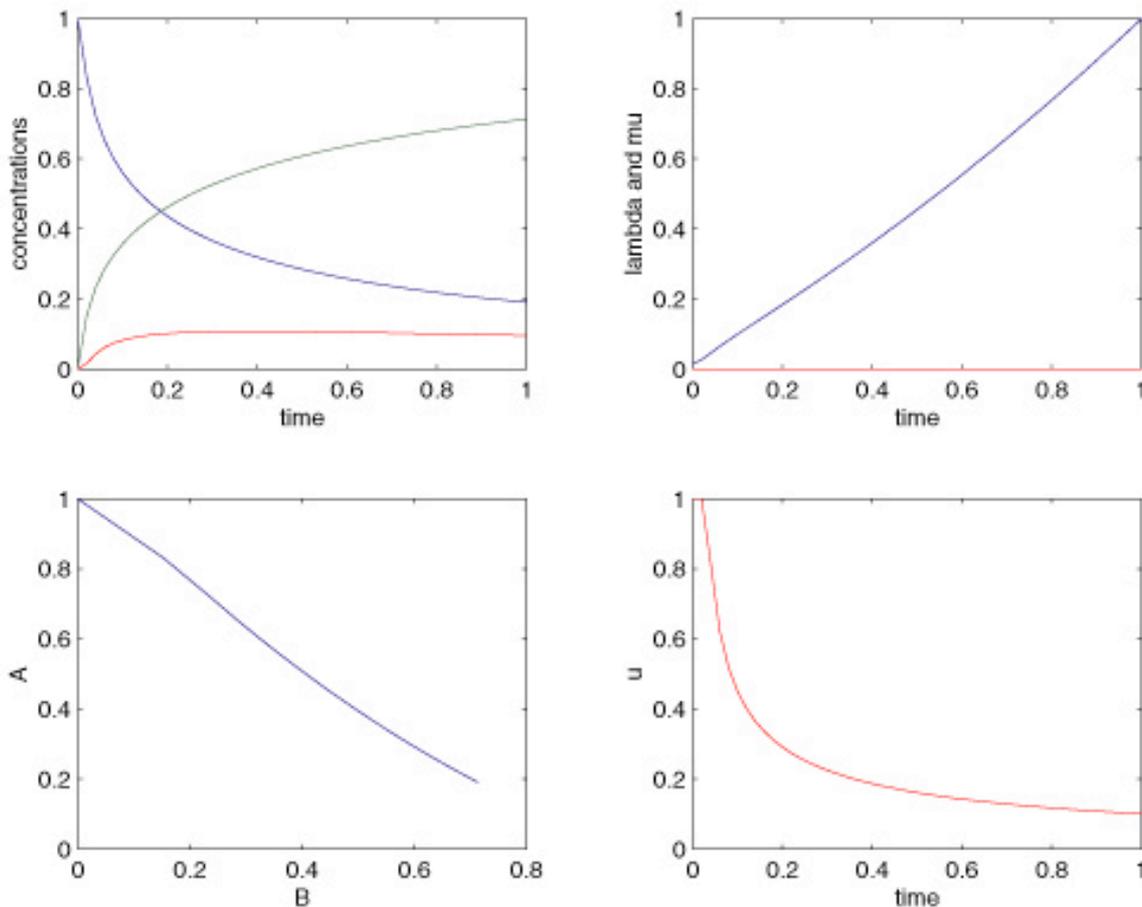


Figure 1: Optimal time sequences of the concentrations A , B , and C (top left: blue, green, red) as well as the conjugate variables λ and μ (top right: blue, red) for activation energies $E=(1, 2, 2, 1)$ and collision factors $s=(0.1, 0.1, 0.1, 0.1)$. The bottom left frame shows the phase relationship, A as function of B , for the optimal path while the bottom right frame shows the time evolution of the control u ("temperature") which produces the optimum yield.

The top two frames in Figure 1 show the optimal time sequences of the concentrations A , B , and C as well as the conjugate variables λ and μ for one such calculation. The bottom left frame shows the time-independent plot of A vs. B . The bottom right frame contains the control u ("temperature") which produces the optimum. The activation energies and collision factors used are $E=(1, 2, 2, 1)$ and $s=(0.1, 0.1, 0.1, 0.1)$, corresponding to B being the most stable component and the equations integrated for a fairly long time. This is

a textbook example with infinite initial temperature, gradually being reduced so as to push the equilibrium toward the desired, energetically favored product. This is an example with infinite maximum useful time Δt_{\min}^* as is the case for any example in which $k_1 = k_4$. In this case it follows that the adjoint variable λ must vanish identically and thus the $H=0$ case forces $dB/dt=0$ contradicting optimality in finite time.

Figure 2 uses $E=(1, 2, 1, 3)$ and $s=(0.7, 0.7, 0.7, 0.7)$. This represents a system where B is more

stable than A and C still more stable. Thus without control the final product would be C. The optimal path calls for a relatively long period at

infinite temperature after which the temperature jumps to a much lower value to prevent loss into C (see frame d).

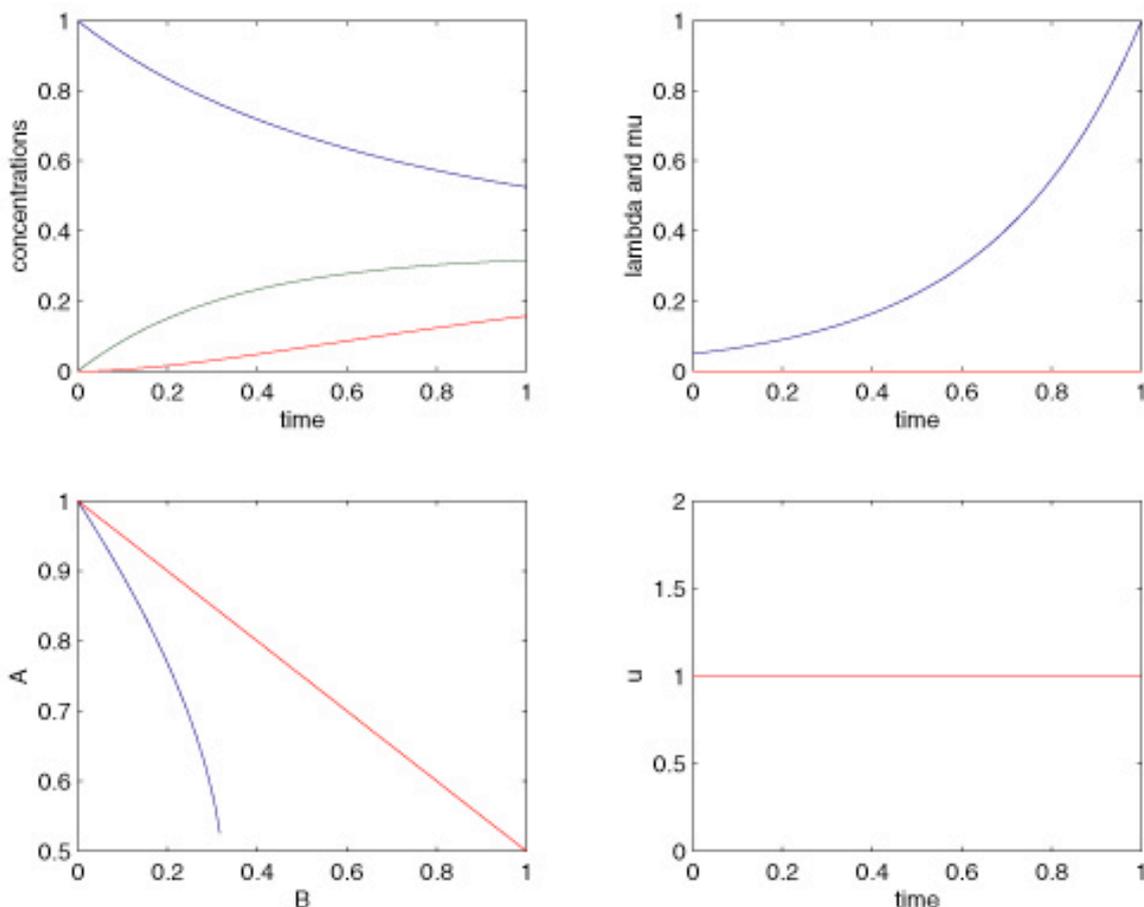


Figure 2: Optimal solution for activation energies $E=(1, 2, 1, 3)$ and collision factors $s=(0.7, 0.7, 0.7, 0.7)$. The nomenclature is the same as in Fig. 1.

Figure 3 with $E=(2, 1, 3, 1)$ and $s=(0.6, 0.6, 0.6, 0.6)$ is the opposite case where A is the most stable substance. Again the optimal path starts with a fairly long period at infinite temperature but then gradually breaks away to lower values as B should be favored over C (lower right frame). In this case for which the optimal temperature moves gradually off down from infinity, we can derive a switching curve shown in frame (c). The cooling begins when the phase trajectory of A vs. B crosses this curve.

A couple of interesting general features should be noted for later discussion.

1. All optimal solutions start with a period of $u=1$ (T infinite) after which the curve

suddenly breaks off to lower temperatures when back reaction begins to be important. This feature is easily proved in general.

2. The final concentrations of A and C are not the lowest / highest seen throughout the calculation. Rather, the reactions seem to have gone a little too far and then return partly to A and B, respectively. This may appear counterproductive, but as long as more B is produced from C than returns to A, the balance for B is still positive.

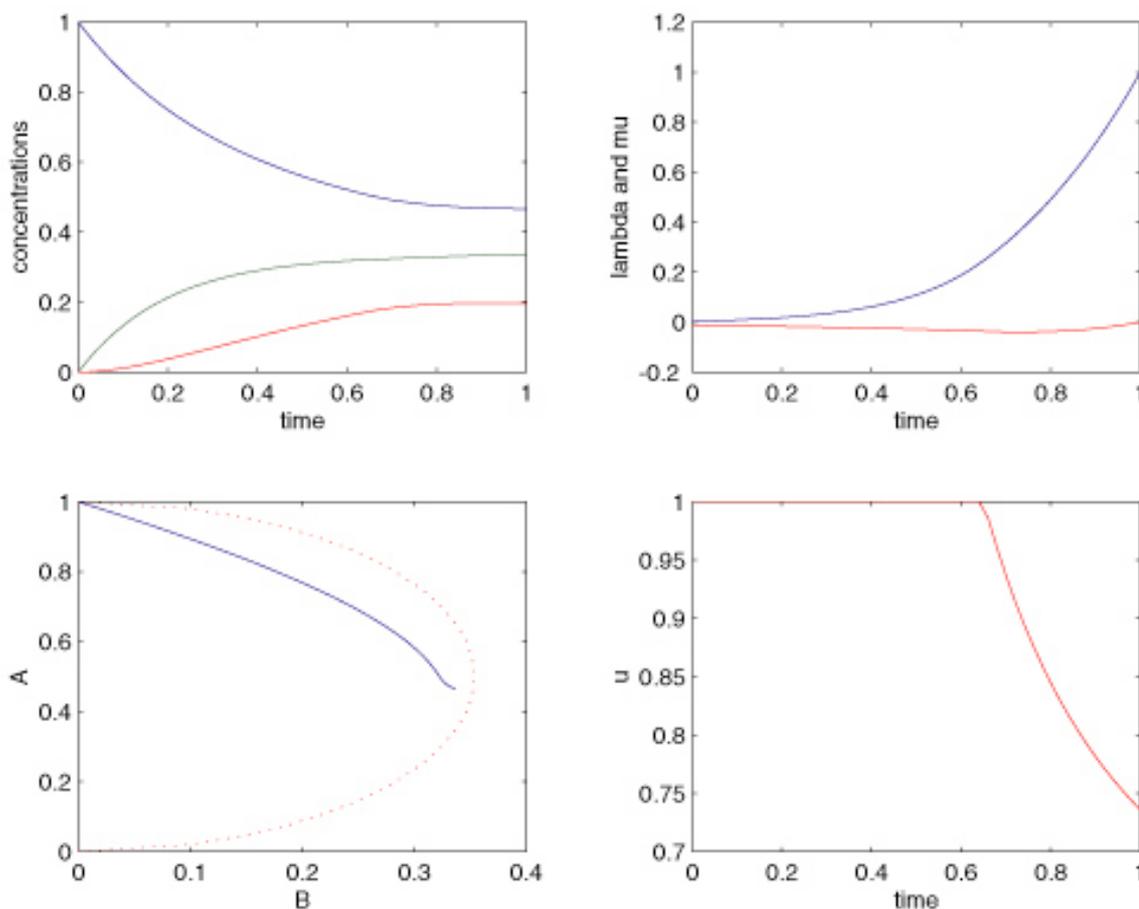


Figure 3: Optimal solution for activation energies $E=(2, 1, 3, 1)$ and collision factors $s=(0.6, 0.6, 0.6, 0.6)$. The nomenclature is the same as in Fig. 1.

CONCLUSIONS

The consecutive chemical reactions $A \rightarrow B \rightarrow C$ have been analyzed both numerically and analytically. Starting with pure A and maximizing the amount of B at the end of the given process duration, we find that all optimal paths start with a segment at infinite temperature ($u=1$). Depending on the energy barriers and the collision factors, the optimal temperature may subsequently switch to either a finite temperature (internal optimum) or zero temperature ($u=0$) if the "maximum useful time" is exceeded. A finite maximum useful time requires that the smallest activation energy in the system be for leaving B. On the other hand, if the two activation energies for leaving B are the largest, the reaction can be driven asymptotically to a yield of 100%.

The time rate of change of B is positive throughout the process. For certain sets of parameter values the production rate dB/dt will approach zero asymptotically for long process durations, for others it will become zero for finite durations. If such an internal zero exists, the optimization becomes indeterminate for longer durations since the required periods of shutting off the reaction can equally well appear anywhere during the process.

The present line of development should be considered part of the continuing quest to answer R. Stephen Berry's original vision of finite-time thermodynamics [12, 13] as the exploration of bounds on what can be achieved in finite time. While the problem discussed in this article does not concern the energetics of the process directly, it is an attempt to find how much of A can be turned into B in a given time and thus belongs to

the realm of finite-time thermodynamics. While we have found several interesting conclusions regarding this simple system, much more remains unanswered.

ACKNOWLEDGEMENTS

We would like to thank Anatoly Tsirlin, and Karl Heinz Hoffmann for many constructive comments. This work is supported by the Danish Natural Science Research Council, contract 21-01-0352. BA wishes to thank San Diego State University for its hospitality.

REFERENCES

- [1] D. Annable, Chem. Eng. Sci. **1**, 145 (1952).
- [2] N. R. Amundson, De Ingenieur No 37, 1 (1955).
- [3] O. Bilous, N. R. Amundson, Chem. Eng. Sci. **5**, 81 115 (1956).
- [4] K. G. Denbigh, Chem. Eng. Sci. **8**, 125 132 (1958).
- [5] F. Horn, U. Troltenier, Chemie Ing. Techn. Nr. 6, 382 393 (1960).
- [6] F. Horn, Z. Elektrochemie **65**, 209 222 (1961).
- [7] F. Horn, Chem. Eng. Sci. **14**, 77 89 (1961).
- [8] R. Aris, Chem. Eng. Sci. **13**, 18 29 (1960).
- [9] R. Aris, Chem. Eng. Sci. **13**, 197 206 (1961).
- [10] R. Aris, Z. Elektrochemie **65**, 229 244 (1961).
- [11] R. Aris, R. S. H. Mah, Chem. Eng. Sci. **19**, 541 553 (1964).
- [12] B. Andresen, P. Salamon, and R.S. Berry, Physics Today, September 1984.
- [13] R.S. Berry, V. Kazakov, S. Sieniutycz, Z. Szwast, and A.M. Tsirlin, Thermodynamic Optimization of Finite-Time Processes, John Wiley & Sons, New York, 2000.
- [14] G. Leitmann, The Calculus of Variations and Optimal Control, Plenum Press, New York, 1981.
- [15] V.F. Krotov, Izvestia of USSR Acad. Nauk, Ser. Techn. Cybern. N5, N6, 1975.