Proceedings of the 17th International Conference
on Computational and Mathematical Methods
in Science and Engineering, CMMSE 2017
4–8 July, 2017.

General Analytical Laws for Metabolic Pathways

L. Bayón[1], P. Fortuny Ayuso[1], J. M. Grau[1], M.M. Ruiz[1] and P.M. Suárez[1]

[1] *Department of Mathematics, University of Oviedo*

emails: bayon@uniovi.es, fortunypedro@uniovi.es, grau@uniovi.es, mruiz@uniovi.es, pedrosr@uniovi.es

Abstract

In this paper a general formulation for the kinetics of multi-step enzymatic reactions is presented. The optimal enzyme and metabolite concentrations are studied for the problem of minimizing the operation time in which the substrate is converted into the product. We give an analytic solution for three different kinetic models for both the unbranched and branched cases. Sufficient conditions for the optimality of the solution are studied. Several examples are presented.

Key words: Optimal Control, Kinetic models
MSC 2000: 49J30, 49M05, 80A30.

1 Introduction

The kinetics of multi-step enzymatic reactions is an ongoing research topic and, in it, the minimization of the operation time in which the substrate is transformed into the product is one of the classical problems which is being currently studied. In it, one measures the optimal profiles of both the enzyme and the metabolite. Our aim in this work is to obtain a general analytic solution for this problem. This way, we avoid the unwieldy numerical solutions, which are always tainted by the specific traits of each particular problem.

In the last years, several results have been presented on the topic. In many of them, an unbranched reaction chain of n irreversible reaction steps is studied (e.g. [1], where an explicit solution for the simplest case, $n = 2$ with linear kinetic model, is given). A mathematical model of an unbranched reaction chain with $n = 3$ and obeying the Michaelis-Menten (MM) kinetic model is used in [2]. In [3], a quasi-analytic solution is found for $n = 3$

 ISBN: 978-84-617-8694-7

and linear kinetics. In [4], the general case of n steps with MM kinetic model is analyzed and quantitative properties are presented, although the authors do not give an explicit analytical solution. In [5], we use a linear kinetic model for the solution of the general case of n steps. Later, in [6], we improve our results with a quasi-analytical solution for the n-steps MM model using the Lambert W-function.

Branched pathways have also been studied, certainly, but only specific cases. In [7], a network inspired in the glycolysis, with the MM model, is considered. The same example is revisited in [8] and a new pathway with two outputs is presented. Other objective functionals may also be considered, like maximizing the productivity of the metabolite [9] or the flux of a particular metabolite [10]. However, in this work we shall minimize the operation time to obtain a specified concentration of the final product.

The numerical methods for the solution of our dynamic optimization problem are usually classified into two groups: direct and indirect methods. Direct methods include complete parametrization [11], multiple shooting [12] or control vector parametrization [13]. In all of them, the basic idea is to transform the original problem into a non-linear programming problem by discretizing and approximating the control and the state variables.

On the other hand, the indirect methods solve the optimization problem using Pontryagin's Minimum Principle (PMP) taking into account the necessary optimality conditions. In this paper, the problem is stated as an Optimal Control Problem (OCP) and using PMP [14] we obtain the solution. Even more (and this is unusual in the literature), we shall also study the sufficient conditions to obtain an optimum. We also remark that we allow the possibility of using three different kinetic models in the same example. Finally, we consider not only an unbranched metabolic pathway but also a branched scheme. We obtain general, model-independent laws, for the first time.

The paper is organized as follows. In Section 2 we present the statement of the problem. The general laws of the optimal solution are obtained in Section 3, and a new kinetic model, the power law, is also presented. Section 4 contains a study on the verification of the sufficient conditions. Then we generalize the problem to the branched case with a statement valid for any graph satisfying some specific conditions. In Section 6 we present a model based on glycolysis which we aim to study numerically.

2 Statement of the problem

For the sake of simplicity we start with the simplest case of an unbranched metabolic pathway composed of n irreversible reaction steps converting substrate x_1 into product p. The value $x_1(t)$ is the substrate concentration at time t, $p(t)$ the concentration of the final product, $x_i(t), (i = 2, \ldots, n)$ the concentration of the intermediate compounds, and $u_i(t)$ $(i = 1, \ldots, n)$ the concentration of the enzyme catalyzing the i-th reaction (see Fig. 1).

 ISBN: 978-84-617-8694-7

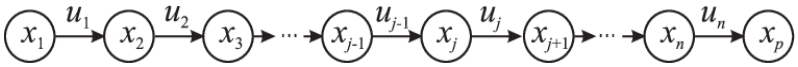

Fig. 1. Unbranched scheme.

Once we have fully studied this case, we shall perform the generalization to the branched one. The rate of the i-th reaction, $v_i(x_i(t), u_i(t))$ is linear in the enzyme concentrations, u_i:

$$v_i(x_i(t), u_i(t)) = w_i(x_i(t)) \cdot u_i(t) \tag{1}$$

The following are frequently used kinetic models:

$$\begin{aligned} w_i(x_i) &= k_i x_i && \text{(Mass action)} \\ w_i(x_i) &= \frac{k_i x_i}{K_i + x_i} && \text{(Michaelis-Menten)} \\ w_i(x_i) &= k_i x_i^c && \text{(Power law)} \end{aligned} \tag{2}$$

The dynamical model for the pathway is given by the law of conservation of mass:

$$\dot{x}_i(t) = v_{i-1}(x_{i-1}(t), u_{i-1}(t)) - v_i(x_i(t), u_i(t)); \ (i = 1, \ldots, n). \tag{3}$$

The objective is to transform x_1 into p as fast as possible; we denote t_f the final time. We assume an exhaustible initial substrate, x_1, and imposing $p(t_f) = C_f$ $(0 < C_f < 1)$, we obtain:

$$x_1(t_f) + x_2(t_f) + \ldots + x_n(t_f) = 1 - C_f \tag{4}$$

so that the optimization problem may thus be stated as the control problem (Pr):

$$\begin{aligned} &\text{(Pr):} && \tau_{C_f} = \min_{u_1, \ldots u_n} \int_0^{t_f} dt = \min_{u_1, \ldots u_n} t_f \\ &\text{subject to:} && (3), (4) \\ & && u_1 \geq 0, \ldots u_n \geq 0; \ u_1 + \ldots + u_n \leq 1 \end{aligned} \tag{5}$$

3 Optimal Solution

In two previous papers, we used PMP to obtain the solution to (Pr) for the mass action model [5] and for the MM model [6]. When the control appears linearly, as is the case for the problem under consideration, the control switches between its upper and lower bounds at discrete instants: the optimal control is said to be a *bang-bang type control* and those instants are called the *switching times*. The general form of the solution can be described

 ISBN: 978-84-617-8694-7

as follows: there exist n switching times, as many as enzymes, so that the optimal i-enzyme profile is proved to be of bang-band type and satisfies:

$$u_i(t) = \begin{cases} 1 & \text{for} \quad t \in [t_{i-1}, t_i) \\ 0 & \text{for} \quad t \notin [t_{i-1}, t_i) \end{cases} \; ; \; i = 1, \ldots, n \tag{6}$$

where $\{t_0, t_1, t_2, ..., t_n\}$ are the switching times, with $t_0 = 0$ and $t_n = t_f$.

We shall denote by $x_{ji}(t)$ (for $i, j = 1, \ldots, n$) the optimal j-th metabolite concentration, in the i-th interval $[t_{i-1}, t_i]$. The optimal solution of the complete system can be described on each interval, knowing that on the i-th interval, $[t_{i-1}, t_i]$ (for $i = 2, \ldots, n-1$), there are 4 laws governing the metabolite concentrations:

(a) Metabolites before the i-th remain at a constant value given by:

$$x_{ji}(t) = x_{jj}(t_j) \text{ for } j = 1, \ldots, i-1 \tag{7}$$

(b) The i-th metabolite follows a law given by a function depending on: the parameters of the model, the previous switching time, and the value of the i-th metabolite on the previous interval:

$$x_{ii}(t) = f(x_{ii-1}(t_{i-1}), t_{i-1}, t) \tag{8}$$

(c) The $i+1$-th metabolite follows a law obtained from the one of the previous metabolite as follows:

$$x_{i+1i}(t) = x_{ii-1}(t_{i-1}) - x_{ii}(t) \tag{9}$$

(d) Metabolites from the $i+2$-th on have not been activated yet, so that their value is zero:

$$x_{ji}(t) = 0 \text{ for } j = i+2, \ldots, n \tag{10}$$

A schematic idea is shown in Fig. 2.

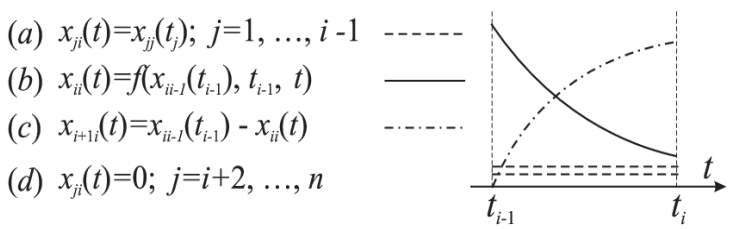

Fig. 2. Optimal concentration laws for the metabolites.

On the first interval ($i = 1$), letting $x_{10}(t_0) = 1$, only Laws (b), (c) and (d) apply; on the last-but-one ($i = n-1$), only Laws (a), (b) and (c) apply; whereas on the last one ($i = n$), only Laws (a) and (b).

 ISBN: 978-84-617-8694-7

Notice that the formulas above are general and they only depend on the kinetic model (Law (b)). In [5] we obtained the law for the mass action model (i.e. the linear one), getting:

$$x_{ii}(t) = f(x_{ii-1}(t_{i-1}), t_{i-1}, t) = x_{ii-1}(t_{i-1}) \exp(-k_i(t - t_{i-1})) \tag{11}$$

Elsewhere, in [6], we obtained the law for the Michaelis-Menten model, which gives:

$$\begin{aligned} x_{ii}(t) &= f(x_{ii-1}(t_{i-1}), t_{i-1}, t) & (12)\\ &= K_{mi} W\left(\frac{x_{ii-1}(t_{i-1})}{K_{mi}} \exp\left(\frac{x_{ii-1}(t_{i-1})}{K_{mi}}\right) \exp\left(-\frac{k_i}{K_{mi}}(t - t_{i-1})\right)\right) & (13)\end{aligned}$$

where W is the Lambert W-function.

We present, for the first time, in this paper, the expression of Law (b) for the kinetic model given by the power law. The expression is now:

$$\begin{aligned} x_{ii}(t) &= f(x_{ii-1}(t_{i-1}), t_{i-1}, t) & (14)\\ &= \left[(x_{ii-1}(t_{i-1}))^{1-c} - (1-c)k_i(t - t_{i-1})\right]^{(1-c)^{-1}} & (15)\end{aligned}$$

The method for computing this optimal solution is analogous to the one we shown in our previous papers, and we refer the reader to them (mainly Appendix 1 in [6], where it is given in detail).

The idea is to define the Hamiltonian $H(x_1, \ldots, x_n, u_1, \ldots, u_n, \lambda_1 \ldots, \lambda_n, t)$ associated to the problem (Pr):

$$\begin{aligned} H &= 1 + \sum_{i=1}^{n} \lambda_i \dot{x}_i(t) & (16)\\ &= 1 + \sum_{i=1}^{n} \lambda_i \left[w_{i-1}(x_{i-1}(t)) \cdot u_{i-1}(t) - w_i(x_i(t)) \cdot u_i(t)\right] & (17)\end{aligned}$$

and compute the optimum values for x_i and u_i applying the necessary conditions in PMP on that Hamiltonian. In this case:

$$\begin{aligned} &\text{(i)} \quad \dot{\lambda}_i(t) = -\frac{\partial H}{\partial x_i}; \lambda_i(t_f) = 0,\ i = 1, ..., n\\ &\text{(ii)} \quad \min_{u_1, \ldots u_n} H\\ &\text{(iii)} \quad \dot{x}_i(t) = w_{i-1}(x_{i-1}(t)) \cdot u_{i-1}(t) - w_i(x_i(t)) \cdot u_i(t);\ x_i(0) = x_{i0},\ i = 1, ..., n\\ &\text{(iv)} \quad H(x_1, \ldots, x_n, u_1, \ldots, u_n, , \lambda_1 \ldots, \lambda_n, t_f) = 0\end{aligned} \tag{18}$$

Our system is autonomous, so that $H_t \equiv 0 \Rightarrow H(t) = cte$. This condition together with (iv) implies that $H(t) = 0$.

 ISBN: 978-84-617-8694-7

We obtain the optimal solution constructively by intervals, starting at $t = 0$ and concatenating the results. Once these values are computed, one still needs to calculate the switching times $t_1, t_2, \ldots, t_{n-1}$ and the operation time t_f. To this end, we use the restriction (4) and define the augmented functional:

$$L(t_1, t_2, \ldots, t_{n-1}, t_f, \beta) = t_f + \beta(x_{1n}(t_f) + x_{2n}(t_f) + \cdots + x_{nn}(t_f) - C_f) \tag{19}$$

where the values of the concentrations $x_{1n}(t_f) = x_{1n}(t_1)$, $x_{2n}(t_f) = x_{2n}(t_2)$, $\ldots, x_{nn}(t_f)$ are given, and where the unknowns $t_1, t_2, \ldots, t_{n-1}$ and t_f appear. Then we solve the non-linear system:

$$\frac{\partial L}{\partial t_1} = 0; \frac{\partial L}{\partial t_2} = 0; \ldots; \frac{\partial L}{\partial t_{n-1}} = 0; \frac{\partial L}{\partial t_f} = 0; \frac{\partial L}{\partial \beta} = 0 \tag{20}$$

Once the optimal values of the switching times $t_1, t_2, \ldots, t_{n-1}$ and t_f are obtained numerically, the remaining values of the solution are immediately obtained analytically using the closed-form formulas (a) to (d), and the problem is completely solved.

Notice that, remarkably, the laws given above allow the simultaneous consideration of several models when studying the pathway: one just needs to use the appropriate function $f(x_{ii-1}(t_{i-1}), t_{i-1}, t)$ in Law (b).

4 Sufficient Conditions

Considering the optimal control problem, with $x(t)$ and $u(t)$ denoting n-dimensional vectors:

$$\min_{u(t)} J = \int_0^T F(x(t), u(t), t)dt + B[T, x(T)] \tag{21}$$

$$\dot{x}(t) = f(x(t), u(t), t);\ x(0) = x_0 \tag{22}$$

$$u(t) \in U(t),\ 0 \le t \le T \tag{23}$$

we can guarantee the sufficient conditions for the existence of an optimal solution using Arrow's Theorem [14] (this sufficiency is not considered in the previous works [4], [5], [6]).

In our case:

$$\min_{u(t)} J = \min_{u_1, \ldots u_n} \int_0^{t_f} dt \tag{24}$$

$$\dot{x}_i(t) = w_{i-1}(x_{i-1}(t)) \cdot u_{i-1}(t) - w_i(x_i(t)) \cdot u_i(t), \quad i = 1, \ldots, n \tag{25}$$

$$u_1 \ge 0, \ldots, u_n \ge 0;\ u_1 + \ldots + u_n \le 1 \tag{26}$$

Upon minimizing $H(x, u, \lambda, t)$ in $u \in U(t)$ one obtains a function $u = u^0(x, \lambda, t)$ from which $H^0(x, u^0(x, \lambda, t), \lambda, t)$ can be computed. In each case, the equations of the model (Mass action, Michaelis-Menten and Power law) provide the sufficient conditions in one way or another.

 ISBN: 978-84-617-8694-7

5 Generalization

We present a generalization for branched pathways. The cases we cover are pathways whose graph of *temporal dependencies* satisfies a specific property (which essentially says that the i-th metabolite follows temporally the j-th one if and only if vertex j is joined with vertex i by a directed path (see [15] for the elementary notions).

These graphs represent the fact that the synthesis of one enzyme requires the degradation of *all the previous ones*; this is the case, for example, of the glycolysis, see [8]. Following that paper, we assume that the optimal profile follows a pattern matching the topology of the pathway, reflecting the fact that the enzymes are activated sequentially. Notice that the type of graphs we deal with have a strictly upper-triangular adjacency matrix (and more conditions, but this one is easy to verify).

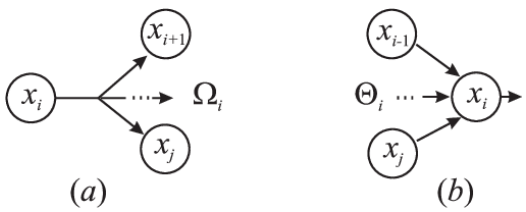

Fig. 3. Branched pathways.

With the same notation as above, the optimal solution of these branched systems can be described on the i-th interval $[t_{i-1}, t_i]$ ($i = 2, \ldots, n-1$) with these 4 laws:

- (a') For the metabolites before the $i-$th one:

$$x_{ji}(t) = x_{jj}(t_j) \text{ for } j = 1, \ldots, i-1 \tag{27}$$

- (b') For the i-th metabolite:

$$x_{ii}(t) = f(x_{ii-1}(t_{i-1}), t_{i-1}, t) \tag{28}$$

- (c') All the metabolites $j \in \Omega_i$ follow the same law:

$$x_{ji}(t) = x_{ji-1}(t_{i-1}) + x_{ii-1}(t_{i-1}) - x_{ii}(t) \tag{29}$$

- (d') The metabolites j-th such that $i < \min \Theta_j$, have not been activated yet:

$$x_{ji}(t) = 0 \tag{30}$$

The only differences with the unbranched system appear in (c') and (d').

 ISBN: 978-84-617-8694-7

6 Numerical Examples

We provide several simulations which illustrate the general formulations above. We shall use a test example already studied by several authors and inspired by the upper part of glycolysis. The original problem was stated by Bartl et al. in [7] and was also considered in [8] with a new formulation incorporating the enzyme dynamics.

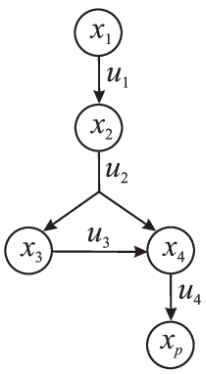

Fig. 4. Glycolysis inspired network.

The pathway (Fig. 4) consists of four enzymatic reactions with one branch. Recall that x_1 corresponds to the substrate, x_2, x_3 and x_4 are the intermediate metabolites and x_p represents the product. The alternative route of the glycolysis is represented by u_2 (corresponding to the enzyme aldolase) metabolizing the intermediate x_2 to x_3 and x_4.

The single aim is to minimize the time needed to transform the substrate x_1 into a fixed amount (90%) of product x_p, (i.e. $C_f = 0.9$). We assume unbuffered or exhaustible substrate x_1 (i.e. the substrate is consumed during the process) and enzymes are assumed to become activated instantaneously (just-in-time activation). The following initial conditions at $t = 0$: $x_0 = [1, 0, 0, 0, 0]^T$; $u_0 = [0, 0, 0, 0]^T$ are imposed. Metabolites and enzymes are expressed in concentration units and time in seconds.

Additionally, restrictions on enzyme concentrations and their total amount are introduced in a normalized form as:

$$u_1 \geq 0, \ldots u_n \geq 0; \; u_1 + \ldots + u_n \leq 1 \tag{31}$$

This is in agreement with the assumption that the cell can only allocate a certain amount of protein to a pathway, and with experimental observations [16] in Escherichia coli.

The classical theoretical studies ([7], [8]) are based on the Michaelis–Menten kinetics:

$$w_i(x_i) = \frac{k_i x_i}{K_i + x_i} \tag{32}$$

with unity rate constants for k_i and K_i: $k_i = 1(s^{-1})$, $K_i = 1(mM)$. However, we also give results for the other two laws we have studied: mass action and the power law.

 ISBN: 978-84-617-8694-7

References

[1] E. Klipp, R. Heinrich, H.G. Holzhutter, *Prediction of temporal gene expression. Metabolic optimization by re-distribution of enzyme activities*, Eur. J. Biochem. **269 (22)** (2002) 5406-5413.

[2] A. Zaslaver, A. Mayo, R. Rosenberg, P. Bashkin, H. Sberro, M. Tsalyuk, M. Surette, U. Alon, *Just-in-time transcription program in metabolic pathways*, Nat. Genet. **36(5)** (2004) 486-491.

[3] M. Bartl, P. Li, S. Schuster, *Modelling the optimal timing in metabolic pathway activation-Use of Pontryagin's Maximum Principle and role of the Golden section*, BioSystems **101** (2010) 67-77.

[4] D. Oyarzun, B. Ingalls, R. Middleton, D. Kalamatianos, *Sequential activation of metabolic pathways: a dynamic optimization approach*, Bull. Math. Biol. **71(8)** (2009) 1851-1872.

[5] L. Bayon, J.A. Otero, M.M. Ruiz, P.M. Suarez, C. Tasis, *Sensitivity analysis of a linear unbranched chemical process with n steps*, J. Math. Chem. **53**(**3**) (2015) 925-940.

[6] L. Bayon, J.A. Otero, P.M. Suarez, C. Tasis, *Solving linear unbranched pathways with Michaelis-Menten kinetics using the Lambert W-Function*, J. Math. Chem. **54(7)** (2016) 1351-1369 .

[7] M. Bartl, M. Kotzing, C. Kaleta, S. Schuster, P. Li, *Just-in-time activation of a glycolysis inspired metabolic network - solution with a dynamic optimization approach*, Proc. 55nd Int. Sci. Colloq. (2010) 217-222.

[8] G. de Hijas-Liste, E. Klipp, E. Balsa-Canto, J. Banga, *Global dynamic optimization approach to predict activation in metabolic pathways*, BMC Syst. Biol. **8(1)** (2014) 1-15.

[9] A.F. Villaverde, S. Bongard, K. Mauch, E. Balsa-Canto, J.R. Banga, *Metabolic engineering with multi-objective optimization of kinetic models*, J. Biotech. **222** (2016) 1-8.

[10] G. Xu, L. Wang, *An improved geometric programming approach for optimization of biochemical systems*, J.Appl. Math. (2014) 1-10.

[11] L.T. Biegler, A.M. Cervantes, A. Watcher, *Advances in simultaneous strategies for dynamic process optimization*, Chem. Eng. Sci. **57(4)** (2002) 575-593.

 ISBN: 978-84-617-8694-7

[12] H.G. Bock, K.J. Plitt, *A multiple shooting algorithm for direct solution of optimal control problems*, In Proceedings 9th IFAC World Congress, New York: Pergamon Press (1984) 242-247.

[13] V.S. Vassiliadis, R.W.H. Sargent, C.C. Pantelides, *Solution of a class of multistage dynamic optimization problems. 1. Problems without path constraints*, Ind. Eng. Chem. Res. **33(9)** (1994) 2111-2122.

[14] A. Chiang, *Elements of Dynamic Optimization*, Waveland Press, 2000.

[15] R. Gould, *Graph Theory*, Dover Publications, 2012.

[16] K.M. Slade, R. Baker, M. Chua, N.L. Thompson, G.J. Pielak, *Effects of recombinant protein expression on green fluorescent protein diffusion in Escherichia coli*, Biochemistry, **48(23)** (2009) 5083-5089.

 ISBN: 978-84-617-8694-7