

# Metabolic Control Analysis and Local Controllability of Biochemical Networks

Brian Ingalls

Department of Applied Mathematics  
University of Waterloo  
Waterloo, Ontario Canada N2L 2G1

Christopher Rao

Chemical and Biomolecular Engineering  
University of Illinois at Urbana-Champaign  
Urbana, IL 61801

## Abstract

We explore issues of local controllability and tracking in biochemical networks using the formalism of Metabolic Control Analysis (MCA). We show that the main results of MCA can be interpreted as solutions to simple tracking problems. The introduction of typical constraints demands an extension of these results to deal with more realistic problems.

## Keywords

Metabolic control analysis, Biochemical networks, Controllability

## Introduction

The topological structure of biochemical reaction networks imposes constraints on their behavior. Though reaction rates are typically nonlinear functions of the concentrations of the interacting species, these rates are incorporated into the dynamics in a linear fashion through the stoichiometric matrix. This special structure can be exploited to yield insight into this class of dynamic systems, which includes metabolic, genetic, and signal transduction networks. In particular, an understanding of how the structure of the reaction network constrains the system response can facilitate pharmacological investigation, gene therapy, and metabolic engineering. In each of these tasks the goal is to perturb the system so that the dynamic and steady-state behavior of the network changes to our advantage.

The consequences of stoichiometry for parametric sensitivity analysis are described by the main theorems of Metabolic Control Analysis (MCA) (Heinrich & Schuster 1996). In this note, we will show that by thinking of parameter perturbations as control inputs these theorems can be translated into local

statements about controllability and tracking. This provides a control-theoretic context for the body of work on redesigning metabolic networks found in the MCA literature. This work began in (Kascer & Acerna 1993) which provided a “universal method” for increasing pathway flux. Later, the problem of increasing metabolite concentration was treated in (Small & Kascer 1994). A local solution to the general design problem was provided in (Westerhoff & Kell 1996), using the main results of MCA. In this paper we provide a systems-theoretic repackaging of the results of Westerhoff and Kell. By making an explicit linearization and posing the objective in terms of tracking, we are able to address the problem in familiar control-theoretic terms. In the simplest case, the problems of reaching a specific state (concentration profile) or tracking a specific flux (reaction profile) are answered by direct appeal to the Theorems of MCA. However, when the target objectives are not consistent or the desired behavior cannot be attained due to lack of control authority, the situation is more complex and the results of the Theorems must be generalized.

Extensions to the nonlinear setting are not treated

here, but would follow along the lines of (Kholodenko, Cascante, Hoek & Westerhoff 1998, Kholodenko, Westerhoff, Schwaber & Cascante 2000).

## Background

Consider a biochemical network described by the following set of differential equations

$$\frac{ds}{dt} = Nv(s, p), \quad (1)$$

where the vector  $s$  contains the concentrations of the  $n$  interacting species, the matrix  $N$  describes the stoichiometry, the vector  $v(\cdot)$  describes the rates for the  $m$  chemical reactions comprising the network, and the vector  $p$  is made up of kinetic parameters which may be subject to perturbation.

We follow the standard approach of Reder (Reder 1988) for dealing with structural conservations among the species concentrations. Let  $r$  denote the rank of  $N$ . Reorder the rows of  $N$  so that the first  $r$  rows are linearly independent. The submatrix consisting of the first  $r$  rows will be called  $N_R$ . Introduce the **link matrix**  $L$  so that  $N = LN_R$ . The species which correspond to the rows of  $N_R$  will be referred to as the **independent species**  $s_i$ . Those remaining are the **dependent species**  $s_d$ . The dependence is made explicit by  $s(t) = Ls_i(t) + T$ , where the constant  $T$  is determined by the initial conditions. Thus, the system dynamics can be written in terms of the  $r$  independent species concentrations

$$\frac{ds_i}{dt} = N_R v(Ls_i + T, p).$$

From this reduction, we see that the original description (1) is typically a non-minimal representation (regardless of the choice of inputs and outputs). The dependence among the species concentrations characterizes  $n - r$  uncontrollable modes.

Given a nominal asymptotically stable steady state  $(s^0, p^0)$  of interest, the linearized dynamics about this point are described by the differential equations

$$\frac{dx_i}{dt} = N_R \varepsilon_s L x_i + N_R \varepsilon_p u, \quad (2)$$

where

$$x_i \triangleq s_i - s_i^0, \quad u \triangleq p - p^0,$$

$$\varepsilon_s \triangleq \left. \frac{\partial v}{\partial s} \right|_{(s^0, p^0)}, \quad \varepsilon_p \triangleq \left. \frac{\partial v}{\partial p} \right|_{(s^0, p^0)}.$$

Linearizing the reaction rate vector as an output of the system, define

$$y \triangleq \varepsilon_s x + \varepsilon_p u \approx v - v^0,$$

where  $x = Lx_i$ . At this stage, a control theorist might ask fundamental questions regarding controllability by considering the pair  $(N\varepsilon_s L, N\varepsilon_p)$ . Here we extend the basic issue of controllability by observing that only states that can be reached and maintained are of biological interest. Thus, rather than controllability, we address issues of tracking by characterizing the concentration and reaction rate profiles that can be reached and maintained by a constant input.

In addressing steady state behavior we must address another constraint imposed by the stoichiometry, namely that a reaction rate vector is consistent with steady state behavior if and only if it lies in the nullspace of  $N$ . Thus, at steady state, the  $m$  reaction rates are restricted to lie in an  $m - r$  dimensional space. This fact, which has been exploited in a number of contexts by Pálsson and colleagues (Schilling & Pálsson 1998), can be formalized as follows (Heinrich & Schuster 1996). Rearranging the columns of  $N$  so that the last  $r$  are linearly independent, the reaction rates  $v$  can be partitioned into independent and dependent reactions as  $v = [v_i^T, v_d^T]^T$ . The dependence is made explicit by choosing a matrix  $K$  of the form  $K = [I_{m-r} \ K_0]^T$  such that the columns of  $K$  span the nullspace of  $N$ . Then at steady state, the reaction rate vector satisfies  $v = K v_i$  so that  $v_d = K_0 v_i$ . This dependence can be seen as analogous to the constraint on the species concentrations. Indeed, both follow from the fact that the stoichiometry matrix is not full rank. In regards to the tracking problem, these both represent constraints on the reference signals which might be tracked. Of the  $n + m$ -dimensional space spanned by the species-reaction output pair, there are only  $m$  dimensions which can be explored by the system.

## Tracking

The local tracking problem can be specified in terms of the linearized system (2) as follows: determine an

input  $u$  that achieves the target  $(x^*, y^*)$ . We can formulate this problem as the solution to the matrix equation

$$\begin{bmatrix} 0 \\ x^* \\ y^* \end{bmatrix} = \begin{bmatrix} N_R \varepsilon_s & N_R \varepsilon_p \\ I & 0 \\ \varepsilon_s & \varepsilon_p \end{bmatrix} \begin{bmatrix} x^s \\ u^s \end{bmatrix}. \quad (3)$$

We begin by considering the simplest case when the target can be achieved.

### Full Authority, No Constraint Violation

Suppose that there are  $m$  independent inputs such that  $\varepsilon_p$  is invertible. Moreover, we will consider the case in which the given target does not violate the stoichiometric constraints. This is equivalent to supposing that the target is specified in terms of the independent variables  $(s_i, v_i)$ .

We first consider the problem of achieving the state target. This problem is described by the matrix equation

$$\begin{bmatrix} 0 \\ x_i^* \end{bmatrix} = \begin{bmatrix} N_R \varepsilon_s L & N_R \varepsilon_p \\ I & 0 \end{bmatrix} \begin{bmatrix} x_i^s \\ u^s \end{bmatrix},$$

which is equivalent to

$$x_i^* = x_i^s = (N_R \varepsilon_s L)^{-1} N_R \varepsilon_p u^s.$$

(Note that the invertibility of the system Jacobian is a consequence of stability of the nominal steady state.) In MCA the operator  $(N_R \varepsilon_s L)^{-1} N_R$  is referred to as the **unscaled concentration control coefficient**  $C^s$ .

The stoichiometric structure of the system allows a decomposition of the inputs  $u^s$  which satisfy this equation. Specifically, for any  $u^s$ , we can write uniquely

$$\varepsilon_p u^s = \varepsilon_s L \pi_1 + K \pi_2,$$

from which we recognize that the state  $x_i^*$  corresponding to  $u^s$  depends on  $\pi_1$  but not on  $\pi_2$ . This fact is made explicit in the concentration Summation and Connectivity Theorems of MCA, which in this context can be described by the following statements:

**Summation:** If  $\varepsilon_p u = K \pi_2$ , then  $x_i^* = 0$ .

**Connectivity:** If  $\varepsilon_p u = \varepsilon_s L \pi_1$ , then  $x_i^* = \pi_1$ .

In particular, to reach the state  $x^*$ , we need to apply an input of the form

$$u = \varepsilon_p^{-1} (\varepsilon_s L x_i^* + K \pi_2)$$

for any choice of  $\pi_2$ .

Likewise, if we consider only the problem of achieving a desired flux, then the problem reduces to the matrix equation

$$\begin{bmatrix} 0 \\ y_i^* \end{bmatrix} = \begin{bmatrix} N_R \varepsilon_s L & N_R \varepsilon_p \\ \varepsilon_s^i L & \varepsilon_p^i \end{bmatrix} \begin{bmatrix} x_i^s \\ u^s \end{bmatrix},$$

which has solution

$$y_i^* = (-\varepsilon_s^i L (N_R \varepsilon_s L)^{-1} N_R \varepsilon_p + \varepsilon_p^i) u^s.$$

The operator  $(-\varepsilon_s L (N_R \varepsilon_s L)^{-1} N_R + I)$  is the **unscaled flux control coefficient**  $C^v$  of MCA. We note that only the projection of  $\varepsilon_p u^s$  onto the nullspace of  $N$  has an impact on the flux. Specifically, if  $\varepsilon_p u^s$  is orthogonal to the nullspace of  $N$ , then  $y_i^* = 0$ . Again, this can be expressed formally by the Flux Summation and Connectivity Theorems:

**Summation:** If  $\varepsilon_p u^s = K \pi_2$ , then  $y_i^* = \pi_2$ .

**Connectivity:** If  $\varepsilon_p u^s = \varepsilon_s L \pi_1$ , then  $y_i^* = 0$ .

Again, the choice of input is clear. To reach the rate profile  $y_i^*$  we can use any input of the form

$$u^s = \varepsilon_p^{-1} (\varepsilon_s L \pi_1 + K y_i^*).$$

Treating both problems together, to achieve the target  $(x_i^*, y_i^*)$ , we take

$$u^s = \varepsilon_p^{-1} (\varepsilon_s L x_i^* + K y_i^*).$$

This result was reached by presuming that the control inputs have full authority over the system and the given target does not violate the constraints imposed by stoichiometry. A more realistic situation is one in which one or both of these assumptions is violated. That more complicated case is considered next.

Decomposition of the concentration and flux vectors into independent and dependent components can be seen as providing a minimal realization of the original system (1). In general, minimal realizations are used to ascertain which combinations of states cannot be independently controlled or observed. In the case of biochemical systems, such constraints are imposed by the stoichiometry and can be systematically addressed through the link matrix  $L$  and the kernel  $K$ .

## Unattainable Targets

Questions of tracking under constraints typically come up against issues of attainability. In this case, the target may be unattainable for two reasons. Firstly, the given target may be incompatible with the stoichiometry. Secondly, the system may be under-actuated, so that only a subset of targets can be reached. In either such case, we can identify the “best we can do” by finding a solution that minimizes the Euclidean distance to the specified target

$$\min_{u^s} \|(x^*, y^*) - C\varepsilon_p u^s\|_2,$$

where  $C \triangleq ((C^s)^T, (C^v)^T)^T$ . In this regard, we are solving a standard least-squares problem.

The solution to this target problem is

$$u^s = ((C\varepsilon_p)^T C\varepsilon_p)^{-1} (C\varepsilon_p)^T (x^*, y^*),$$

which results in the steady state

$$(x^s, y^s) = C\varepsilon_p ((C\varepsilon_p)^T C\varepsilon_p)^{-1} (C\varepsilon_p)^T (x^*, y^*).$$

Thus the optimal output is the projection of the target onto the row space of  $C\varepsilon_p$ .

The special cases in which either  $\varepsilon_p$  or  $C$  are invertible are also of interest. If  $\varepsilon_p$  is invertible, then the input has full authority over the independent concentrations and reaction rates. In this case the least-squares solution yields a steady state of

$$(x^s, y^s) = C(C^T C)^{-1} C^T (x^*, y^*),$$

the projection of the target onto the row space of  $C$ . Alternatively, if the target satisfies the stoichiometric constraints it can be specified by the pair  $(s_i, v_i)$ . In this case, the matrix  $C$  is replaced with the invertible  $C_i = ((C^{s_i})^T, (C^{v_i})^T)^T$ . The optimal steady state is then

$$(x_i^s, y_i^s) = C_i \varepsilon_p (\varepsilon_p^T \varepsilon_p)^{-1} \varepsilon_p^T (x^*, y^*),$$

which is the result of projecting onto the row space of  $\varepsilon_p$ . Here the control input provides an insufficient number of degrees of freedom to achieve the desired objective. The resulting approximation is the projection of the target onto the set of attainable profiles.

## Conclusion

This note has highlighted a connection between Metabolic Control Analysis and classical control theory. By posing the tracking problem for linearized models of biochemical systems appropriately, we confirm that the main results of MCA can be seen as providing the solution to this problem in the simplest case. Extension to more realistic problems requires a generalization of the results to provide optimal approximate solutions when given tracking targets are unattainable.

## References

- Heinrich, R. & Schuster, S. (1996), *The Regulation of Cellular Systems*, Chapman and Hall.
- Kascer, H. & Acernoza, L. (1993), ‘A universal method for achieving increases in metabolite production’, *Eur. J. Biochem.* **216**, 361–367.
- Kholodenko, B. N., Cascante, M., Hoek, J. B. & Westerhoff, H. V. (1998), ‘Metabolic design: how to engineer a living cell to desired metabolite concentrations and fluxes’, *Biotechnology and Bioengineering* **59**, 239–247.
- Kholodenko, B. N., Westerhoff, H. V., Schwaber, J. & Cascante, M. (2000), ‘Engineering a living cell to desired metabolite concentrations and fluxes: pathways with multifunctional enzymes’, *Metabolic Engineering* **2**, 1–13.
- Reder, C. (1988), ‘Metabolic control theory: A structural approach.’, *J. Theor. Biol.* **135**, 175–201.
- Schilling, C. H. & Palsson, B. O. (1998), ‘The underlying pathway structure of biochemical reaction networks’, *Proc Natl Acad Sci U S A.* **95**, 4193–4198.
- Small, J. R. & Kascer, H. (1994), ‘A method for increasing the concentration of a specific internal metabolite in steady state systems’, *Eur. J. Biochem.* **226**, 649–656.
- Westerhoff, H. V. & Kell, D. B. (1996), ‘What biotechnologists knew all along ...?’, *J. Theor. Biol.* **182**, 411–420.