NORTHERN ILLINOIS UNIVERSITY

Mathematical Models of Biochemical Switch Networks

A Capstone Submitted to the

University Honors Program

In Partial Fulfillment of the

Requirements of the Baccalaureate Degree

With Honors

Department Of

Mathematical Sciences

By

Cassandra Mohr

DeKalb, Illinois

May 2019

University Honors Program
Capstone Approval Page

Capstone Title (print or type) Mathematical Models of Biochemical Switch Networks

Student Name (print or type) Cassandra Mohr

Faculty Supervisor (print or type) MAYA MINCHEVA

Faculty Approval Signature [signature]

Department of (print or type) Mathematical Sciences

Date of Approval (print or type) 5/3/19

Check if any of the following apply, and please tell us where and how it was published:

☐ Capstone has been published (Journal/Outlet):

☐ Capstone has been submitted for publication (Journal/Outlet):

☐ Capstone has been presented (Conference):

☐ Capstone has been submitted for presentation (Conference):

Completed Honors Capstone projects may be used for student reference purposes, both electronically and in the Honors Capstone Library (CLB 110).

If you would like to opt out and not have this student's completed capstone used for reference purposes, please initial here: _______ (Faculty Supervisor)

HONORS CAPSTONE ABSTRACT

This project centers on mathematical applications to biochemistry. Specifically, the use of a dynamical system to model a special type of biochemical network and determine the effect of initial concentrations on the existence of several constant solutions. Many biochemical networks act as biological switches that are responsible for important biological functions such as cell differentiation and cell death; consequentially, the ability to better predict and manipulate their outcome is of great importance. One particularly insightful and relatively simple form of biochemical mechanism is that of the reversible substrate inhibition reaction. Utilizing basic principles of mathematics and chemistry, it is possible to convert a biochemical network into a system of differential equations; this in turn permits further in-depth analysis of the original chemical reaction in order to determine its projected outcomes. Using the Jacobian and its characteristic polynomial as well as analysis of the system itself, we can gain enhanced comprehension into the effect initial concentration has on the eventual outcome of the overall system of biochemical reactions. Specifically, it is possible to determine what limitations must be imposed on the initial network in order to guarantee fixed points.

Mathematical Models of Biochemical Switch Networks

Cassandra Mohr

Honors Capstone

Saddle-node bifurcations are the basic mechanism for the creation and destruction of fixed points; that is, as a parameter varies, two fixed points can bifurcate from a single one. These fixed points influence the flow of the overall system, determining the general trends and outcomes associated with it. The analysis of the saddle-node bifurcation for a multidimensional system of differential equations can permit greater understanding of the system, allowing us to determine the influence that initial conditions have on the outcome of the system. By analyzing the bifurcations of a biochemical mechanism, we can effectively determine constraints that will result in a manageable, effective reaction[1].

One particularly insightful and relatively simple form of biochemical mechanism is that of the reversible substrate inhibition reaction. The term substrate inhibition refers to the process wherein the substrate of an enzyme reaction inhibits the enzyme's activity. In the simplest of terms, a substrate can be defined as the primary reactant in a chemical reaction; in other words, the substance on which the enzyme reacts. An enzyme, in turn, is a substance which acts as a catalyst to bring about a specific biochemical reaction. The particular biochemical reaction being considered is as follows:

$$A_1 \overset{k_1}{\to} (),$$

$$() \overset{k_2}{\to} A_1,$$

$$A_1 + A_2 \overset{k_3}{\to} A_3,$$

$$A_3 \overset{k_4}{\to} A_2,$$

$$A_1 + A_3 \overset{k_5}{\to} A_4,$$

$$A_4 \overset{k_6}{\to} A_1 + A_3$$

In this system of elementary reactions, A_1 is a substrate, A_2 is an enzyme, and A_3 and A_4 are enzyme-substrate complexes; k_i for each $i \in \{1, . . . ,6\}$ is a positive rate constant[2]. The arrows (→) represent influx and outflux.

[1] Strogatz, S. H. (1994). Nonlinear dynamics and Chaos: With applications to physics, biology, chemistry, and engineering. Reading, Mass.: Addison-Wesley Pub..

[2] Mincheva, M., Roussel, M.: Graph-theoretic methods for the analysis of chemical and biochemical networks. I. Multistability and oscillations in ordinary differential equation models. Mathematical Biology. 61-86 (2007).

Per the law of mass action, there exists a proportionality between the rate of an elementary reaction and the product of the concentrations of the reactants; as such, it is possible to convert the above series of elementary reactions into a series of differential equations[3]. Let the concentrations of A_k be u_k for each $k \in \{1, . . . ,4\}$. Then the system of differential equations derived from our initial reaction is

$$\begin{aligned}
\frac{du_1}{dt} &= -k_1u_1 + k_2 - k_3u_1u_2 - k_5u_1u_3 + k_6u_4 \\
\frac{du_2}{dt} &= -k_3u_1u_2 + k_4u_3 \\
\frac{du_3}{dt} &= k_3u_1u_2 - k_4u_3 - k_5u_1u_3 + k_6u_4 \\
\frac{du_4}{dt} &= k_5u_1u_3 - k_6u_4
\end{aligned} \tag{1}$$

Additionally, we can also derive initial conditions for this system:

$$\begin{aligned}
u_1^o &= u_1(0) \\
u_2^o &= u_2(0) \\
u_3^o &= u_3(0) \\
u_4^o &= u_4(0)
\end{aligned}$$

It is notable that the last three equations- $\frac{du_2}{dt}$, $\frac{du_3}{dt}$, and $\frac{du_4}{dt}$ – are linearly dependent; that is, their sum equals zero. This indicates that

$$u_2 + u_3 + u_4 = u_2(0) + u_3(0) + u_4(0)$$

Since k_i is always positive and u_i are concentrations, it can be shown that $u_i \geq 0$.

Our overall goal is to determine how the initial concentrations- that is, u_k for each $k \in \{1, . . . ,4\}$- affect the overall behavior of the system and specifically, the existence of several fixed points. In particular, we want to determine what constraints are necessary to produce several fixed points; in other words, we desire the production of equilibrium solutions whose overall outcomes are not affected by small disturbances away from them[3]. In order to attain this, we need to further analyze the last non-zero coefficient of (1)'s characteristic polynomial; doing so will permit the development of suitable constraints.

Deriving the Jacobian of (1), we obtain the following matrix:

[3] Strogatz, S. H. (1994). Nonlinear dynamics and Chaos: With applications to physics, biology, chemistry, and engineering. Reading, Mass.: Addison-Wesley Pub..

$$J = \begin{bmatrix} -k_1 - k_3u_2 - k_5u_3 & -k_3u_1 & -k_5u_1 & k_6 \\ -k_3u_2 & -k_3u_1 & k_4 & 0 \\ k_3u_2 - k_5u_3 & k_3u_1 & -k_4 - k_5u_1 & k_6 \\ k_5u_3 & 0 & k_5u_1 & -k_6 \end{bmatrix}$$

Here we assume that J is evaluated at some fixed point. The determinant of this matrix is equal to zero as a result of the linear dependence of rows 2, 3, and 4. We now want to analyze the characteristic polynomial of the above determinant; that is, p(λ)= det(J- λI) = $\lambda^4 + a_1\,\lambda^3 + a_2\,\lambda^2 + a_3\lambda + a_4$. Since a_4 = det(-J), we know that a_4 = 0. As such, det(J- λI) = $\lambda^4 + a_1\,\lambda^3 + a_2\,\lambda^2 + a_3\lambda = \lambda(\lambda^3 + a_1\,\lambda^2 + a_2\,\lambda^1 + a_3)$. Consequentially, we will focus on the value of a_3, since this is the last non-zero coefficient of p(λ). We consider the following matrix:

$$\begin{bmatrix} -k_1 - k_3u_2 - k_5u_3 & -k_3u_1 & -k_5u_1 & k_6 \\ -k_3u_2 & -k_3u_1 & k_4 & 0 \\ k_3u_2 - k_5u_3 & k_3u_1 & -k_4 - k_5u_1 & k_6 \\ k_5u_3 & 0 & k_5u_1 & -k_6 \end{bmatrix} - \begin{bmatrix} \lambda & 0 & 0 & 0 \\ 0 & \lambda & 0 & 0 \\ 0 & 0 & \lambda & 0 \\ 0 & 0 & 0 & \lambda \end{bmatrix}$$

Which can be simplified to:

$$\begin{bmatrix} -k_1 - k_3u_2 - k_5u_3 - \lambda & -k_3u_1 & -k_5u_1 & k_6 \\ -k_3u_2 & -k_3u_1 - \lambda & k_4 & 0 \\ k_3u_2 - k_5u_3 & k_3u_1 & -k_4 - k_5u_1 - \lambda & k_6 \\ k_5u_3 & 0 & k_5u_1 & -k_6 - \lambda \end{bmatrix}$$

Calculating the determinant of this matrix, we obtain the following value as the coefficient of λ:

$$a_3 = k_1k_3k_6u_1 + k_1k_3k_5u_1^2 - k_3k_4k_5u_1u_3 + k_1k_4k_6 + k_3k_4k_6u_2$$

Which is a quadratic equation with respect to u_1:

$$(k_1k_3k_5)u_1^2 + (k_1k_3k_6 - k_3k_4k_5u_3)u_1 + (k_1k_4k_6 + k_3k_4k_6u_2)$$

Setting this equal to zero and employing the quadratic formula to solve for u_1, we obtain:

$$u_1 = \frac{k_3k_4k_5u_3 - k_1k_3k_6 \pm \sqrt{(k_1k_3k_6 - k_3k_4k_5u_3)^2 - 4k_1k_3k_5(k_1k_4k_6 + k_3k_4k_6u_2)}}{2k_1k_3k_5}$$

Let the two roots of u_1 be denoted as u_1^1 and u_1^2. In other words, let

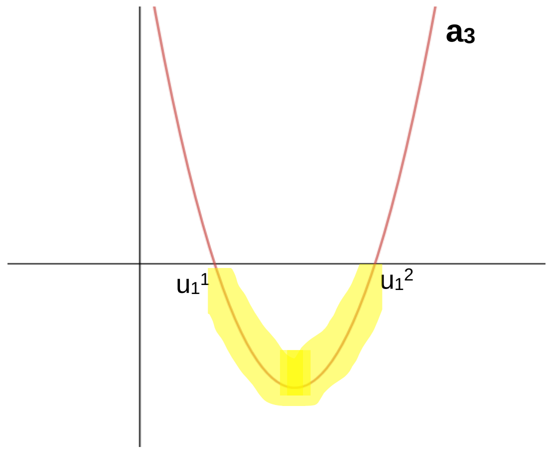
$$u_1^1 = \frac{k_3k_4k_5u_3 - k_1k_3k_6 - \sqrt{(k_1k_3k_6 - k_3k_4k_5u_3)^2 - 4k_1k_3k_5(k_1k_4k_6 + k_3k_4k_6u_2)}}{2k_1k_3k_5}$$

$$u_1^2 = \frac{k_3k_4k_5u_3 - k_1k_3k_6 + \sqrt{(k_1k_3k_6 - k_3k_4k_5u_3)^2 - 4k_1k_3k_5(k_1k_4k_6 + k_3k_4k_6u_2)}}{2k_1k_3k_5}$$

It is vital for u_1^1 and u_2^1 to both be positive so that $a_3 \leq 0$ for some parameter values, thus causing the existence of a saddle-node bifurcation. If they are both positive, then

$$k_1 k_3 k_6 - k_3 k_4 k_5 u_3 < 0$$

We wish to impose constraints that ensure that u_1 or u_2 is real; that is, that any slight change in the initial conditions will damp out over time. In order to gain an increased understanding of the overall system, it can be useful to graphically study the sign of a_3:



In order for $a_3 < 0$ for some parameter values, we need $u_1^1 < u_1 < u_1^2$ (graphically, for u_1 to be in the highlighted region of the graph). This constraint will ensure that any initial u_1 in this domain will have to eventually flow to one of the two nodes; as such, it is impossible for the system to spiral out of control away from these two nodes[4].

As we also want to determine when $a_3 = 0$ will have two real roots, we will now focus in on the discriminant; more specifically, when the discriminant is greater than zero. The contents of the discriminant can as such be expanded and rearranged to generate the following polynomial with respect to u_3:

$$D = (k_3^2 k_4^2 k_5^2)u_3^2 + (-2k_1 k_3^2 k_4 k_5 k_6)u_3 + (k_1^2 k_3^2 k_6^2 - 4k_1^2 k_3 k_4 k_5 k_6 - 4k_1 k_3^2 k_4 k_5 k_6 u_2) \ (*)$$

As we can see, u_1 will have two real roots when the above quadratic equation with respect to u_3 is greater than zero. Consequentially, we need to obtain a greater understanding of (*) from above.

Once again setting this equal to zero and solving for the roots of u_3, we now obtain that the roots of u_3 are equal to:

---

[4] Strogatz, S. H. (1994). Nonlinear dynamics and Chaos: With applications to physics, biology, chemistry, and engineering. Reading, Mass.: Addison-Wesley Pub..

$$\frac{2k_1k_3^2k_4k_5k_6 \pm \sqrt{4k_1^2k_3^4k_4^2k_5^2k_6^2 - 4k_3^2k_4^2k_5^2(k_1^2k_3^2k_6^2 - 4k_1^2k_3k_4k_5k_6 - 4k_1k_3^2k_4k_5k_6u_2)}}{2k_3^2k_4^2k_5^2}$$

This can be further simplified to the following equation:

$$u_3 = \frac{k_1k_3k_6 \pm \sqrt{k_1^2k_3k_4k_5k_6 + k_1k_3^2k_4k_5k_6u_2}}{k_3k_4k_5}$$

Recall now that the purpose of these calculations was to determine when u_1 has two real positive roots. By finding the roots of the discriminant- with respect to u_3- of u_1, we now have a more comprehensive understanding of the sufficient conditions for $a_3 = 0$ to have real roots. Let the two roots of u_3 be denoted as u_3^1 and u_3^2. Then we can say that

$$u_3^1 = \frac{k_1k_3k_6 - \sqrt{k_1^2k_3k_4k_5k_6 + k_1k_3^2k_4k_5k_6u_2}}{k_3k_4k_5}$$

$$u_3^2 = \frac{k_1k_3k_6 + \sqrt{k_1^2k_3k_4k_5k_6 + k_1k_3^2k_4k_5k_6u_2}}{k_3k_4k_5}$$

where $u_3^2 > 0$ and u_3^1 can have either sign. As we wish to focus on when u_3 is positive- and thus the discriminant of u_1 is positive and generates two positive roots- the value of u_3 needs to be constrained so that $0 < u_3 < u_3^1$ or $u_3 > u_3^2$.

From the above calculations, we can conclude that if u_3 is in the interval $I_1 = (0 , u_3^2)$ in the case $u_3^1 < 0$ or if u_3 is in the interval $I_2 = (0 , u_3^1)U(u_3^2 , \infty)$ then the discriminant D of $a_3 = 0$ is positive. Additionally, if $D > 0$, then u_1^1 and u_1^2 are real. If we pick a value $u_3^* = u_3(k_1 , . . . , k_6)$ where u_3 is either in I_1 or I_2, then u_1^1 and u_1^2 are both positive since $\frac{k_1k_4k_6+k_3k_4k_6u_2}{k_1k_3k_5} > 0$.

Thus, for u_1 in (u_1^1 , u_1^2), $a_3 < 0$. Finally, using Proposition 13 applied to a_3, we can claim that a saddle-node bifurcation occurs[5].

Returning now to the linear dependence of $\frac{du_2}{dt}, \frac{du_3}{dt}$, and $\frac{du_4}{dt}$, we wish to analyze the relationship

$$(u_2 + u_3 + u_4) = e_0 , e_0 = u_2(0) + u_3(0) + u_4(0)$$

Setting the original four differential equations equal to zero, we obtain the following system:

---

[5] S. Bosi, D. Desmarchelier (2017) A simple method to study local bifurcations of three and four-dimensional systems : characterizations and economic applications FAERE Working Paper, 2017.16.

$$(1) - k_1u_1 + k_2 - k_3u_1u_2 - k_5u_1u_3 + k_6u_4 = 0$$
$$(2) - k_3u_1u_2 + k_4u_3 = 0$$
$$(3) \quad k_3u_1u_2 - k_4u_3 - k_5u_1u_3 + k_6u_4 = 0$$
$$(4) \quad k_5u_1u_3 - k_6u_4 = 0$$

Adding together (1) and (4) and manipulating the results, we obtain the following:

$$u_1 = \frac{k_2}{k_1 + k_3u_2}$$

As we see that u_1 can be written in terms of u_2, we will fix u_2 in order to simplify the calculations; consequentially, u_2 is a parameter. Plugging in the values obtained for u_1 and u_2 above (both functions of u_2), we subsequently obtain the following functions of u_2 for u_3 and u_4:

$$u_3 = \frac{k_2k_3u_2}{k_1k_4 + k_3k_4u_2}$$

$$u_4 = \frac{k_2^2k_3k_5u_2}{(k_1k_6 + k_3k_6u_2)(k_1k_4 + k_3k_4u_2)}$$

Now, we want to expand $(u_2 + u_3 + u_4)=e_0$ using these new formulas for the variables:

$$u_2 + u_3 + u_4 = e_0 \implies$$

$$u_2 + \frac{k_2k_3u_2}{k_1k_4 + k_3k_4u_2} + \frac{k_2^2k_3k_5u_2}{(k_1k_6 + k_3k_6u_2)(k_1k_4 + k_3k_4u_2)} = e_0 \implies$$

$$(k_3^2k_4k_6)u_2^3 + (2k_1k_3k_4k_6 + k_2k_3^2k_4k_6 - e_0k_3^2k_4k_6)u_2^2 + (k_1^2k_4k_6 + k_1k_2k_3k_6 + k_2^2k_3k_5 - 2e_0k_1k_3k_4k_6)u_2 - e_0k_1^2k_4k_6 = 0 \ (2)$$

As we already know, the coefficients k_j where $j \in \{1, . . . ,6\}$ and u_i where $i \in \{1, . . . ,4\}$ are always positive; additionally, e_0 is positive. This indicates that the coefficient for u_2^3 is positive; concurrently, the coefficient for the free term is negative. If the coefficients for the u_2^2 and u_2^1 terms can be manipulated to be negative and positive respectively, then by Descartes' Rule u_2 has either 1 or 3 positive roots. In order for (2) to have 3 positive real roots, we impose the following constraints:

$$2k_1k_3k_4k_6 + k_2k_3^2k_4k_6 - e_0k_3^2k_4k_6 < 0$$
$$k_1^2k_4k_6 + k_1k_2k_3k_6 + k_2^2k_3k_5 - 2e_0k_1k_3k_4k_6 > 0$$

This implies that

$$\frac{k_1^2 k_4 k_6 + k_1 k_2 k_3 k_6 + k_2^2 k_3 k_5}{2k_1 k_3 k_4 k_6} > e_0 > \frac{2k_1 k_3 k_4 k_6 + k_2 k_3^2 k_4 k_6}{k_3^2 k_4 k_6}$$

If these constraints are met, the overall system may have 3 positive fixed points; as such, small changes in the initial conditions will be damped over time. Likely, 2 of the fixed points will be stable and one unstable. Resultantly, the system will avoid becoming unstable over time; it will not spiral out of control. Imposing such constraints permits us to ensure that the initial biochemical reaction being studied is not unstable.

While we have found constraints that may produce a system with 3 positive fixed points, there is still much room for improvement. Future research into the necessary parameters for stable and unstable fixed points and a resultant overall stable system is necessary and would improve the success of the model. However, the constraints found in this paper serve as a solid basis which enables greater control over the reversible substrate inhibition reaction and has the potential to generate a more ideal and successful reaction network.