Investigation of the *lac* Operon

Diego Wahl[1]

[1]*Department of Physics, Washington University, St. Louis, Missouri 63130*

(Dated: May 8, 2019)

The *lac* operon controls transportive and metabolic processes of lactose in *Escherichia coli*. Explorations of models of this have previously led to a number of different effects and characteristics. This paper focuses on work on analysis of the Yildirim-Mackey model in describing the previously discovered bistability of the system.

INTRODUCTION

Operons, first introduced by Jacob, et. al. (1960)[1], have been a heavily studied concept since their inception. Work has been done on a number of models, including ones to form both repressible (forming a negative feedback)[2] and inducible (positive)[3] systems. More modern models have expanded on previous works to create both stochastic and dynamical models.

In particular, the *lac* operon in *E. Coli* system controls the genes necessary for the use of lactose in the bacterium. In broadest terms, the lactose (*lac*) operon in consists of a cluster of three genes, *lacZ, lacY,* and *lacA*, as well a promoter/operator region (P/O). Preceding the system is a regulator gene *lacI* which is responsible for producing a repressor (R) protein. The repressor binds to O in the absence of allolactose (A) and inhibits RNA polymerase from transcribing structural genes. However when allolactose is present, R cannot bind to the operator region, allowing the RNA polymerase to bind and initiate transcription of mRNA (M). The *lacZ* gene encodes mRNA to eventually produce an enzyme named β-galactosidase (B), and translation of *lacY* forms mRNA responsible for production of another enzyme, permease (P). lacA also forms its own complex, however it is believed not to be of relevence and thus disregarded in most texts. Further, in the presence of external lactose (L_e), the permease transports L_e into the cell to form intracellular lactose (L), which is then broken down into allolactose and eventually glucose and galactose by B. A schematic overview of the system can be seen in Figure (1).

This paper focuses on modern techniques and models, taking special interest in applying numerical methods to approximate solutions.

THE YILDIRIM-MACKEY MODEL

The so-called Yildirim-Mackey model, introduced originally by N. Yildirim, and M. C. Mackey (2003)[4], has been a major focus of research with regards to the *lac* operon model since its inception. The model is governed by the dynamics of 5 concentrations, β-galactosidase B, mRNA M, intracelluar lactose L, allolactose A, and permease P. The model, as presented by Yildirim and Mackey is as follows:

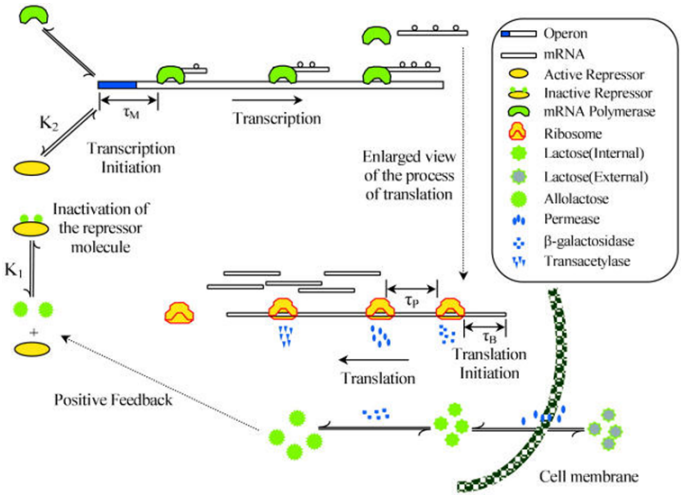

FIG. 1: Schematic of the *lac* operon system as given by N. Yildirim, M. C. Mackey (2003)[4]

$$\frac{dB}{dt}(t) = \alpha_B e^{-\mu\tau_B} M_{\tau_B} - \tilde{\gamma_B} B \tag{1}$$

$$\frac{dM}{dt}(t) = \alpha_M \frac{1 + K_1(e^{-\mu\tau_M} A_{\tau_M})^n}{K + K_1(e^{-\mu\tau_M} A_{\tau_M})^n} + \Gamma_0 - \tilde{\gamma_M} M \tag{2}$$

$$\frac{dL}{dt}(t) = \alpha_L \frac{[P]L_e}{K_{L_e} + L_e} - \beta_{L_1} \frac{P \cdot L}{K_{L_1} + L} - \beta_{L_2} \frac{B \cdot L}{K_{L_2} + L} - \tilde{\gamma_L} L \tag{3}$$

$$\frac{dA}{dt}(t) = \alpha_A \frac{B \cdot L}{K_L + L} - \beta_A \frac{B \cdot A}{K_A + A} - \tilde{\gamma_A} A \tag{4}$$

$$\frac{dP}{dt}(t) = \alpha_P e^{-\mu(\tau_B + \tau_P)} M_{\tau_B + \tau_P} - \tilde{\gamma_P} P \tag{5}$$

Experimental values of parameters, along with justifications of terms in the 5 variable model are included in cited text.

The Reduced, 3 Variable Model

Unfortunately, analysis the 5-variable Yildirim-Mackey model proves numerically too difficult for the purposes of this paper. However, the system can be reduced by making a handful of assumptions:

1. Constant permease P concentration.

2. Across the membrane lactose is in a quasisteady state, and hence internal and external (L_e) lactose are directly proportional.

By assuming as such, we can eliminate both lactose and permease dynamics from our system[5], and we can instead describe our dynamics by the equations (6), (7), and (9).

β-galactosidase dynamics will be governed by:

$$\frac{dB}{dt}(t) = \alpha_B e^{-\mu\tau_B} M_{\tau_B} - \tilde{\gamma_B} B \tag{6}$$

In the production term, one can note that production of β-galactosidase is proportional to mRNA concentration. The $e^{-\mu\tau_B} M_{\tau_B}$ term enforces a time delay due to the time of translation from mRna to β-galactosidase. Note that we define M_{τ_B}:=M(t-τ_B) to express the time delay of τ_B.

Meanwhile the subtractive term accounts for both degradation and diffusion of β-galactosidase. This arises from the definition $\tilde{\gamma_i} := \gamma_i + \mu, i \in (M, B, L, A, P)$, where $gamma_i$ is the degradation parameter, and μ the diffusive one.

mRNA dynamics will be governed by:

$$\frac{dM}{dt}(t) = \alpha_M \frac{1 + K_1(e^{-\mu\tau_M} A_{\tau_M})^n}{K + K_1(e^{-\mu\tau_M} A_{\tau_M})^n} - \tilde{\gamma_M} M \tag{7}$$

The production term of mRNA encapsulates that it is proportional to the fraction of free operators, i.e.;

$$\frac{O}{O_{Tot}} = \frac{1 + K_1(e^{-\mu\tau_M} A_{\tau_M})^n}{K + K_1(e^{-\mu\tau_M} A_{\tau_M})^n} \tag{8}$$

Lastly, allolactose dynamics will be governed by:

$$\frac{dA}{dt}(t) = \alpha_A \frac{B \cdot L}{K_L + L} - \beta_A \frac{B \cdot A}{K_A + A} - \tilde{\gamma_A} A \tag{9}$$

The first term of the right-hand side of 9 represents the production of allolactose from lactose by β-galactosidase, and the second term represents the conversion from allolactose to glucose and galactose by β-galactosidase.

Under these assumptions, the system reduces to such an extent that analysis may continue.

Steady State Analysis

By applying the normal steady state conditions to our system, we can look for fixed points. Allowing $\frac{dA}{dt} = \frac{dM}{dt} = \frac{dB}{dt} = 0$ and solving for B in the allolactose (A) equation yields, after substitution back into equations (7), (6):

$$\frac{1 + K_1(e^{-\mu\tau_M} A_{\tau_M})^n}{K + K_1(e^{-\mu\tau_M} A_{\tau_M})^n} = \frac{\tilde{\gamma_M}\tilde{\gamma_B}\tilde{\gamma_A} e^{\mu\tau_B}}{\alpha_M \alpha_B \alpha_A} \cdot \frac{A}{\frac{L}{K_L + L} - \frac{\beta_A}{\alpha_A}\frac{A}{K_A + A}} \tag{10}$$

By allowing:

$$f(A_{ss}) = \frac{1 + K_1(e^{-\mu\tau_M} A_{\tau_M})^n}{K + K_1(e^{-\mu\tau_M} A_{\tau_M})^n} \tag{11}$$

$$h(L) = \frac{L}{K_L + L} \tag{12}$$

$$g(A_{ss}) = \frac{A_{ss}}{K_A + A_{ss}} \tag{13}$$

$$\chi = \frac{\tilde{\gamma_M}\tilde{\gamma_B}\tilde{\gamma_A} e^{\mu\tau_B}}{\alpha_M \alpha_B \alpha_A} \tag{14}$$

then, substitution into 10 reduces the function to:

$$f(A_{ss}) = \chi \frac{A_{ss}}{h - \frac{\beta_A}{\alpha_A} g} \tag{15}$$

In order to explore the steady states, this paper chooses to first analyse them graphically. To do so, one plots the left-hand side (LHS), which corresponds to equation (11). Then, solve the right-hand side (RHS) for variable values of L, and plot alongside the LHS. Where the RHS curves intersect the LHS curve indicates a fixed point of the system with those parameters.

Of importance to note in Figure (2) is the red line corresponding to L=40μM, namely that it crosses the LHS curve thrice, implying that there are 3 fixed points for that value of L. In fact, expanding the graphical solution to further times, one can find that the system transitions from bi-stability to tri-stability at some value between L=35 and L=45.

To further explore this behavior, (11) can be rearranged to solve for L=L(A_{ss}) as

$$L(A_{ss}) = \frac{K_L \cdot m}{1 - m} \tag{16}$$

Where

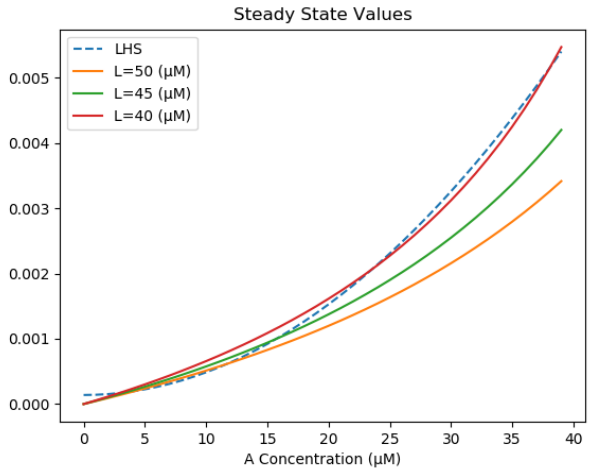

FIG. 2: Steady states as a function of allolactose concentration, found for varying values of L (μM). Note the number of times each solid line (RHS) crosses the dotted line (LHS). Sample code can be found in *354SS.py*

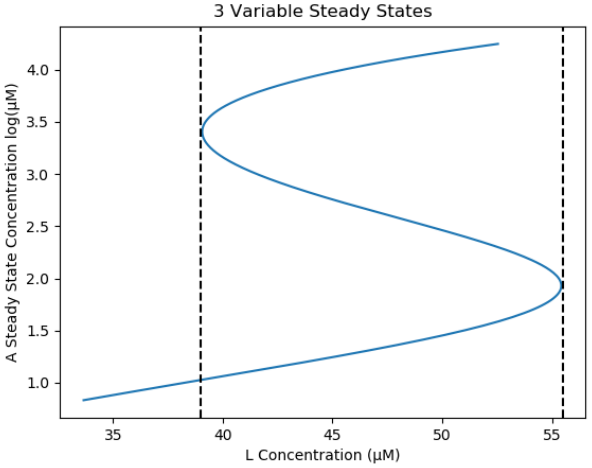

FIG. 3: Steady States in (L,A_{ss}) space. Vertical lines at x=38 and x=55.5 bind the region of 3 fixed points. Sample code can be found in *354AvL.py*

$$m = \frac{\chi \cdot A + f\frac{\beta_A}{\alpha_A} g}{f} \tag{17}$$

While keeping all other values constant, one can create figures in (L,A_{ss}) space such as Figure (3). From the figure it can be discerned that within a range (roughly bound by the dotted vertical lines) of L values the lac operon system does indeed display bi-stability, with the stable fixed points of A_{ss} where the dotted lines touch the filled line in Figure (2).

Further, attempts were made to analyze the dynamics of this deterministic model by solving the series of delay differential equation (DDE), to reproduce graphs such as given in Fig. (3) in Yildirim, Mackey (2003)[4] which dynamically simulate the varying concentrations. However, the numerical techniques involved proved too difficult, and the attempt unsuccessful. The attempted code is attached in *354ddeAttempt.py*.

CONCLUSION/REMARKS

Regarding the time evolution of the reduced, 3 variable model; attempts were made in a number of different coding languages (namely: MATLAB via dde23, Python via JiTCDDE, and R via ddesolve) to resolve the system via a numerical of the DDE. This however, as noted before, went unsuccessfully. The difficulty may have in part stemmed from the sparse documentation and examples of such systems to apply to the *lac* operon system. In the merit of further academic research, the author suggests further modifications to the model as needed. For instance, removing time-delay and thus forming a system of ODEs may produce dynamics still worthy of study.

Secondly, a more rigorous analysis can still be done regarding the steady states of our reduced model. This paper did not fully explore the different regimes of types of fixed points, specifically regimes with two fixed points.

In all, though dissatisfied at having failed to numerical solved the reduced Yildirim-Mackey model of lac operons, the author remains intrigued by the system. The still young and unexplored biological system clearly exhibits (though through some effort) both regions of bistability and regions without. In the future, the author would like to explore stochastic simulations (such as through the Gillepsie algorithm) of the model, rather than deterministic, such as in the work of Carrier and Keasling[6].

CODE

Sample code used to create figures is provided below via the author's personal GitHub repository.

354SS.py: https://git.io/fjc1X
354AvL.py: https://git.io/fjc1i
354ddeAttempt.py: https://git.io/fjc11

[1] D. P. C. S. Jacob, F. and J. M. 1960, C. R. Acad. Sci. **250**, 1727–1729 (1960).
[2] J. S. Griffith, Journal of Theoretical Biology **20**, 202 (1968a).
[3] J. S. Griffith, Journal of Theoretical Biology **20**, 209 (1968b).
[4] N. Yildirim and M. C. Mackey., Biophysical Journal **84**, 2841 (2003).

[5] N. Yildirim, M. Santillán, D. Horike, and M. C. Mackey, Chaos: An Interdisciplinary Journal of Nonlinear Science **14**, 279 (2004), https://doi.org/10.1063/1.1689451, URL https://doi.org/10.1063/1.1689451.

[6] T. Carrier and J. D. Keasling, Journal of theoretical biology **201 1**, 25 (1999).