ANALYTIC SOLUTIONS FOR DIFFUSION ON PATH GRAPHS AND ITS APPLICATION TO THE MODELING OF THE EVOLUTION OF ELECTRICALLY INDISCERNIBLE CONFORMATIONAL STATES OF LYSENIN

by
K. Summer Ware

A thesis

submitted in partial fulfillment

of the requirements for the degree of

Master of Science in Mathematics

Boise State University

December 2020

BOISE STATE UNIVERSITY GRADUATE COLLEGE

DEFENSE COMMITTEE AND FINAL READING APPROVALS

of the thesis submitted by

K. Summer Ware

Thesis Title: Analytic Solutions for Diffusion on Path Graphs and its Application to the Modeling of the Evolution of Electrically Indiscernible Conformational States of Lysenin

Date of Final Oral Examination: 23 October 2020

The following individuals read and discussed the **thesis** submitted by student K. Summer Ware, and they evaluated the student's presentation and response to questions during the final oral examination. They found that the student passed the final oral examination.

Uwe Kaiser, Ph.D.	Chair, Supervisory Committee
Daniel Fologea, Ph.D.	Member, Supervisory Committee
Michal Kopera, Ph.D	Member, Supervisory Committee

The final reading approval of the thesis was granted by Uwe Kaiser Ph.D., Chair of the Supervisory Committee. The thesis was approved by the Graduate College.

DEDICATION

To Tyler.

ACKNOWLEDGMENTS

I want to extend my genuine gratitude to my advisors for guiding me through this process. I also want to thank my husband for supporting me unconditionally through it all.

AUTOBIOGRAPHICAL SKETCH

I am a twenty-five year old Master's student at Boise State University. I was awarded two bachelor's degrees from the same institution for physics and mathematics in 2018. I have a loving husband, two geese, two ducks, and too many houseplants.

ABSTRACT

Memory is traditionally thought of as a biological function of the brain. In recent years, however, researchers have found that some stimuli-responsive molecules exhibit memory-like behavior manifested as history-dependent hysteresis in response to external excitations. One example is lysenin, a pore-forming toxin found naturally in the coelomic fluid of the common earthworm *Eisenia fetida*. When reconstituted into a bilayer lipid membrane, this unassuming toxin undergoes conformational changes in response to applied voltages. However, lysenin is able to "remember" past history by adjusting conformational state based not only on the amplitude of the stimulus but also on its previous its conformational state. The current model is a simple two-state Markov description which may not describe a system with memory. In this respect, this thesis aims to provide a more accurate description of this toxin's memory and response to external stimuli by applying a more rigorous mathematical approach. The traditional setting for investigating the conformational changes of voltage-responsive channel proteins is based on analyzing the ionic currents recorded through one or many channels in response to applied voltage stimuli. However, this approach provides only indirect evidence of the conformational state of the channel, i.e open (conducting) or closed (non-conducting). Although very useful, this setting is seriously limited by the inability of electrical measurements to discern between electrically identical yet conformational different open or closed states. The litera-

ture that inspired this thesis topic consider models of diffusion on a path-graph with one open state and an arbitrary number of closed states. The mathematics typically begins with approximations from a continuous model. In this thesis we study the analytic solution of the system of linear homogeneous differential equations which are probability vectors describing the diffusion process; this involves exponential theory of weighted Laplacian graphs. Since the Laplacian matrix of the path graph is well studied, we have access to both eigenvectors and eigenvalues in terms of roots of unity making for a succinct solution. We find that polynomial weights model the hysteresis successfully.

TABLE OF CONTENTS

LIST OF FIGURES

x

xi

CHAPTER 1:

INTRODUCTION

Hysteresis is a fundamental feature of a physical system, which occurs when a system evolves upon disturbance by following a certain pathway but returns to its original state by adopting a different, distinct pathway. Such features are currently exploited for technological purposes, and memories based on electric and magnetic hysteresis have been largely used in computers, phones, and consumer electronics. Memory is one of the most important applications of hysteresis; however, some biological systems also present hysteretic behavior, but the physiological implications of such behavior have yet to be deciphered. Although biological hysteresis could be a protective measure against strong and transient stimulation or noise, its association with memory provides opportunities for endowing cells with molecular memory. This is of upmost importance for physiology since memory is typically thought of as a function of the brain. Nonetheless, it has been demonstrated that simple, unicellular organisms may develop memory of their own and use it as a survival skill. How is this memory function realized in such simple organisms? Based on experimental evidence, scientists identified ion channels as presenting a hysteretic behavior in response to physical stimuli. Ion channels are protein molecules that assemble into transmembrane structures providing conducting pathways for ions and molecules across cell membranes. They are regulated, which means that they open and close (gate) in response to physical

and chemical stimuli. Voltage-gated ion channels, which open and close in response to transmembrane voltages are responsible for establishing essential electrochemical gradients across cell membranes; their opening and closing is an integral part of the correct functionality of our brain and muscles! Nonetheless, little is known about their hysteretic behavior that may lead to molecular memory. These molecules cannot do algebra or recite poetry, but they are able to "remember" the environmental conditions they were in while holding a specific conformational state. For example in the paper "Voltage Activation and Hysteresis of the Non-selective Voltage-dependent Channel in the Intact Human Red Blood Cell" the authors are able to apply periodic voltages to human red blood cells and see that the non-selective cation channels displayed memory (3). Other ion channels also present hysteresis in conductance, but it manifests only at very short time scales.

Lysenin is a weakly-selective (does not distinguish between ion types) pore-forming toxin made up of 297 amino acids, which is found in the coelomic fluid of the common earthworm (see Figure 1.1 (10)). This toxin has hysteretic behavior that lasts for long periods of time (on the scale of 14 hours or more) in a laboratory setting which makes the other examples pale in comparison. This gives researchers plenty of time to experiment with lysenin before the bilayer membrane breaks down. Lysenin is also water soluble and relatively inexpensive making it an even more viable experimental tool. In one afternoon a new experiment could be started and concluded which is almost unheard of in the study of cellular biology! We do not have the means to watch what happens when a lysenin pore gates. Since we have no physical proof of what is going on, the best we can do to explain the hysteresis is through mathematical modeling.

Typically when it comes to theoretically modeling a physical system we will begin from the ground up by working with well-studied mathematical formulas and theorems before discussing the physics. Here though, the system is so complicated (due to its biological nature) such that empirical methods need to be utilized to study it. This new approach to a model will be a hybrid of mathematics, in the form of solving a system of homogeneous, linear, differential equations, as well as empirical evidence. To solve these equations a path graph and its Laplacian matrix will be employed. These tools are well understood in graph theory and are a natural introduction to the current classical model that is used by those studying lysenin. In one of the papers that inspired this research the author began with a continuous model of diffusion, and through approximations, worked his way down to a discrete case (12). The method we have chosen to employ in this paper is diagonalization of Laplacian path graph matrices and the application of the exponential map. The point in using this approach is to find the closed-form spectrum of the path graph and then from there finding the closed-form eigenvectors and achieve hysteresis in conductance, which yields the conclusion of this thesis.

Figure 1.1   An image of *Eisenia fetida* or the common red wiggler worm.(6)

CHAPTER 2:

BIOPHSYICS

Lysenin self-inserts into artificial and natural lipid membranes containing sphin-gomyelin. The pore formation process is intricate and implies outside assembly of a nonamer (having nine-parts) pre-pore, after which the pores is completed by de-ploying the beta-sheets into the target membrane, see Figure 2.1.
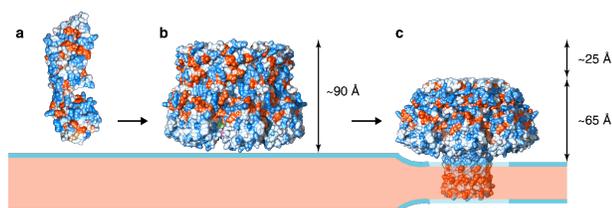
Figure 2.1   (a) A lysenin monomer. (b) The nine-part pre-pore is formed. (c) Lysenin inserted into the bilayer membrane. The pore's opening is $\sim$ 3nm in diameter. (7)

The experimental setup for investigating the response to external voltage stimuli is simple, see Figure 2.2. It involves a bilayer membrane chamber that consists of two insulating reservoirs filled with an electrolyte solution and separated by a thin PTFE film that contains a small central hole of $\sim 70\mu$m diameter. A tiny amount of a lipid mixture in an organic solvent is deposited over the central hole and a lipid membrane is produced by using a combination of painting and folding procedures. The connections with the electrophysiology amplifier is realized via two Ag/AgCl electrodes inserted

into the ionic solutions in the two reservoirs. The entire setup is placed on a floating table inside a Faraday cage to minimize vibrations and electrical noise. The analog signal is digitized, visualized in real time, and recorded for further analysis. The lipid membrane formation is monitored by real-time measurements of conductance (for leakage) and capacitance (for assessment of the number of lipid layers). Once a stable bilayer lipid membrane is formed, lysenin channel reconstitution is realized by adding the lysenin monomer to the grounded reservoir while the active electrode is biased by negative voltages to prevent voltage-induced gating. Channel insertion is observed as step-wise, small and uniform changes in the ionic currents; completion of the insertion process is achieved in approximately 1 hour.
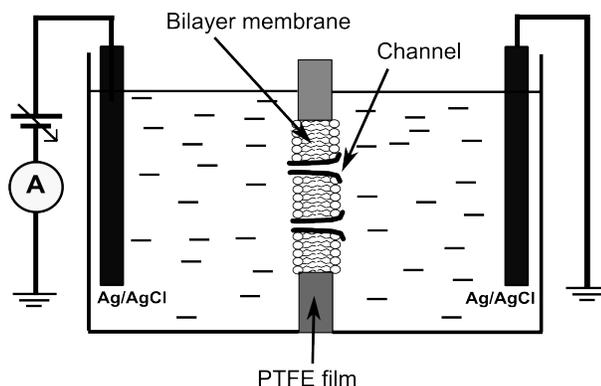
Figure 2.2   A diagram of the experimental set up (9).

There is no direct confirmation of the crystal structure of the domain sensor so we can only surmise from data what it may look and behave like. There is evidence of a flexible, hinge-like structure that is necessary for pore formation. The theory is that the illusive voltage domain sensor is on the tip of this hinge which may be located at the "head" of the pore in the α region; the "mushroom cap" (see figure 2.1). Since the pores respond to pH changes and the ionic strength of the solution
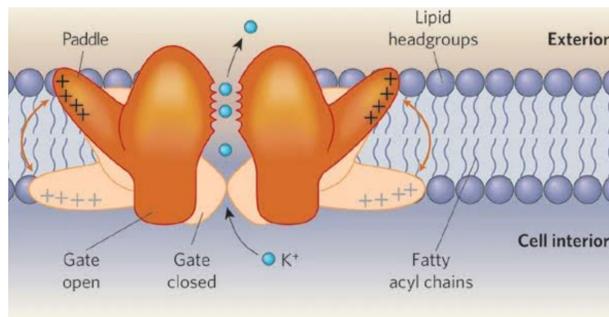
Figure 2.3 A potassium ion channel which shows evidence to be structurally very similar to lysenin (11).

they are in, it is reasonable to conclude that at rest (when the channels are in a conducting state) the voltage domain sensor is fully exposed to the solution. When the channels are closed at high electric potential it is thought that the domain sensor enters the bilayer membrane changing the confirmation of the channel by squeezing the opening of the pore shut. Imagine that each of the nine parts of the lysenin channel has a thumb sticking up above the membrane's surface, and when a large positive potential is present, that thumb forces itself into the hydrophobic part of the bilayer membrane thereby puckering the pore shut. This hypothesis is based on the functionality of a potassium channel which has this gating mechanism and behaves similarly to lysenin with respect to voltage-induced gating (see Figure 2.3). The channels, upon reopening, follow a similar path along the hysteresis independent of ionic concentration (5). Since ions cannot enter a membrane while attached to the domain sensor, it is believed that the sensor is screened of all ions lending credence to the idea that the voltage domain sensor is attached to the hinge of the pore. The current two-state model (open/closed) largely adopted for describing the response of channels to voltage and depicted in Figure 2.4 is based on a reversible transition between the two states with the closing and opening rates denoted as k_+ and k_-,

respectively.
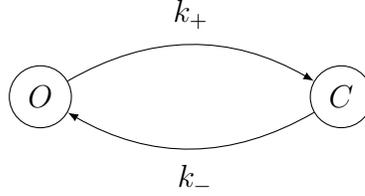
Figure 2.4   The classical two-state process for lysenin.

At any given applied voltage, the system achieves a steady state. This is a dynamic equilibrium since transitions between states still occur but no changes in occupancy are observed at reasonable large time scales. According to the transition state theory, the probability of channels to be in the open state (P_{Open}) at equilibrium is described by a Boltzmann distribution:

$$P_{Open} = \frac{1}{1 + e^{\frac{-\Delta E + qV}{kT}}} \tag{2.1}$$

where ΔE is the difference in energy between the closed and open states, V is the applied voltage, k is the Boltzmann constant, T is temperature in Kelvins, and q is the gating charge of the voltage domain sensor interacting with the electric field and responsible for gate movement.

To observe the current we employ Ohm's Law, $I = V/R$ (where V is the applied voltage and G is macroscopic conductance) and obtain:

$$I = VgNP_{Open} \tag{2.2}$$

where N is the total number of channels in the membrane and g is the conductance (I/V) of a single channel. Eqs. 2.1 and 2.2 may be combined to provide a relationship

more suitable for experiments that imply measuring ionic currents:

$$I = \frac{gVN}{1 + e^{\frac{-\Delta E + qV}{kT}}} + G_L V \tag{2.3}$$

In the above relationship, G_L is the leakage conductance, which manifests as small ionic currents that are recorded through channels in the closed state (when P_{open} approaches zero). Also, when thousands of channels are reconstituted in a target membrane, a few of them may remain in the open state even at high voltages, and this current is accounted for as leakage. Also, each individual channel presents a very small leakage current even in the closed state, and the contribution from all these currents is included in the leakage term.

One may easily observe that Eq. 2.3 describes a reversible response to applied voltages, and that the P_{Open} value is dependent only on voltage. Sometimes a channel may be partially blocked and cannot shut properly leading to leakage current. This means that the equilibrium constant K_e, which is equal to $\frac{k_+}{k_-}$ does not change. This invariance leads to a typical response of ion channels to voltage, i.e. the open probability depends only on voltage and no hysteretic behavior is observed. However, lysenin channels present an atypical response to periodic voltages and do not obey the rule of invariant equilibrium constant. This is exemplified in Figure 5.1, which shows the record of the ionic currents through lysenin channels in response to periodic voltages.

When the transmembrane voltage is linearly ramped up, lysenin channels are all in the open state (conducting) at bias potentials under ~ 12 mV. In the I-V plot, this is observed as a linear increase of the macroscopic ionic current with voltage, indicative of absence of gating. As the voltage is increased, the macroscopic currents
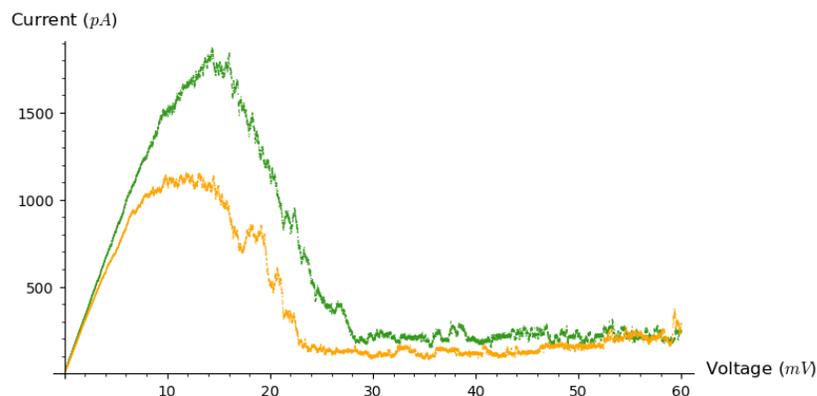
Figure 2.5   A typical hysteresis for a population of lysenin channels. The green curve shows how ionic current decreases as voltage is increased in-dicating that the channels are gating. The yellow curve indicates how the current increases as the voltage decreases indicating that the channels are in a conducting state.

deviate from linearity, reach a peak value, and decrease monotonically with the applied voltage. This is indicative of channel closing (voltage-induced gating), and it is explained by a summary analysis of Eqs. 2.1 and 2.3. As the voltage increases, the exponent increases significantly hence the open probability and ionic currents will approach a near zero value. However, for a markovian process a completely reversed behavior and otherwise identical I-V would be expected when the closed channels are subjected to linearly descending voltage ramps. Nonetheless, Figure 5.1 shows a different reactivation (re-opening) pathway, indicative that the channels do not reopen in exactly the same fashion they closed. This unexpected hysteresis in conductance may reside in the slow inactivation of the channels, i.e. the slow response to applied voltages. The closing of any channel is a statistic process, and no channel closes instantly when a proper transmembrane voltage is applied. Lysenin channels close slowly in response to voltages, and the characteristic relaxation time is ~ 10 s (8). Theoretical models indicate that slow closing may be at the origin of ion channels'

hysteresis, but this must vanish if the voltage varies so slowly such that each state may be considered steady. For the case of lysenin, the hysteresis does not vanish even when the period of the signal is as large as 14 hours, which is much larger that the characteristic relaxation time (8). Therefore, lysenin channels presents hysteresis in conductance and memory. One may easily observe from the IV plot that for any voltage between 8 mV and 28 mV the macroscopic currents depend not only on the applied voltage but on the history. The macroscopic currents are significantly larger if the channels were previously in the open state, while previously closed channels exhibit smaller ionic currents at the same applied voltage. This dependency of the past history is a potential source of molecular memory. The traditional model of gating explicitly disregard any dependency of the states, leading to memory-less channels. It is clear that such models are not adequate for describing the behavior of lysenin channels.

To fill this gap in the knowledge, the biophysical models of gating may be improved by either considering that the channel undergo multiple yet electrically indiscernible closed states, or that the gating process is better characterized as fractal diffusion in the conformational space. Both models account for variable equilibrium constants, which is a great departure from the classical model of gating. The model adopted for analysis in this thesis considers that the channels are able to adopt multiple closed states that may be arranged in a linear kinetics scheme. When a closing voltage is applied, the channel will explore the closed conformational states without presenting changes in the ionic currents (which will be near zero for all the closed states). However, for a linear kinetics scheme the apparent equilibrium constant depends on the state the channels are in; consequently, as the time passes the equilibrium constant

varies, which may lead to hysteresis.

There is also some experimental evidence in favor of having multiple electrically indiscernible closed states. For a typical step experiment it takes 3 or 4 exponential terms to even closely model the closing of the channels in response to step voltages. This signifies that there is something about the closing of the channels that is not being seen by our current two-state model, see Figure 2.6. So to expand upon the foundation of the model described, we will need to further the mathematical understanding of the biophysics. In the following sections we will work to build up our model to include more closed states as well as time-dependent rate constants.
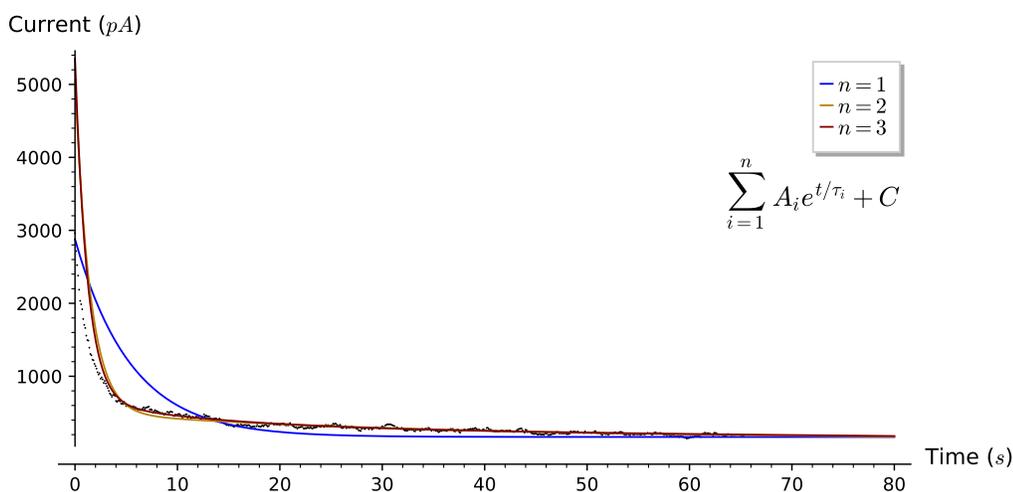
Figure 2.6   Three or four exponential terms in the sum on the graph are needed to analyze data from a typical experiment lending credence to the possibility of more undetectable closed states.

It would be good at this point to mention some of the works that have inspired this paper. "Diffusion models of ion-channel gating and the origin of power-law distributions from single-channel recording" used a continuous model of closed-states with a single open state for their ion channel. They then modeled the transition between

each state as a differential equation where the solution would be the probability $p_n(t)$ of being in that particular state at a given time t. In the paper they concluded that a fractal diffusion (in the form of a polynomial dependent on t and number of channels N) was the best model to explain how their ion channels gated (12). Another paper that influenced this thesis was "Computing rates of Markov models of voltage-gated ion channels by inverting partial differential equations governing the probability density functions of the conducting and non-conducting states" which explored having multiple closed states with exponential rate constants (1). Finally, Reference (8) titled "Bi-stability, hysteresis, and memory of voltage-gated lysenin channels" proves that individual lysenin channels have hysteretic behavior.

Definition 3.0.1. Weighted Digraph or Weighted Directed Graph

A *weighted digraph* is a graph whose vertices or edges have been assigned weights and whose edges have associated direction.

Typically in the literature one will take a graph like Figure 3.1 and represent it mathematically as a system of differential equations (12). These equations, when solved, will tell us the probability P_{C_i} of being in a given closed state C_i and the probability P_{Open} of being in the open state P_{Open} at a given time or voltage. We derive these equations directly from Figure 3.1 by using directions and weights of edges to and from each vertex.

$$\frac{dP_{Open}}{dt} = (P_{C_1})(k_{1-}) - (P_{Open})(k_{1+})$$

$$\frac{dP_{C_1}}{dt} = (P_{Open})(k_{1+}) - (P_{C_1})(k_{1-}) + (P_{C_2})(k_{2-}) - (P_{C_1})(k_{2+})$$

$$\vdots$$

$$\frac{dP_{C_i}}{dt} = (P_{C_{(i-1)}})(k_{i+}) - (P_{C_i})(k_{i-}) + (P_{C_{(i+1)}})(k_{(i+1)-}) - (P_{C_i})(k_{(i+1)+})$$

$$\vdots$$

$$\frac{dP_{C_n}}{dt} = (P_{C_{(n-1)}})(k_{n+}) - (P_{C_n})(k_{n-})$$

With

$$P = (P_{Open}, P_{C_1}, P_{C_2}, ..., P_{C_{n-1}}, P_{C_n})^T$$

and

$$
L_w = \begin{pmatrix}
-k_{1+} & k_{1-} & & \cdots & & 0 \\
k_{1+} & -(k_{2+} + k_{1-}) & k_{2-} & \cdots & & 0 \\
0 & k_{2+} & -(k_{3+} + k_{2-}) & \cdots & & 0 \\
\vdots & & & \ddots & & \vdots \\
0 & \cdots & & & -(k_{(n-1)-} + k_{(n)+}) & k_{(n)-} \\
0 & \cdots & & 0 & k_{(n)+} & -k_{(n)-}
\end{pmatrix}
$$

the system of equations becomes $\frac{dP}{dt} = L_w \cdot P$ where P is the vector of probabilities describing the dynamics of our system above, immediately derived from Figure 3.1.

Lemma 3.0.1. *If an $n \times n$ matrix A commutes with the matrix $\frac{d}{dt} A$, then*

$$
\frac{d}{dt} A^n = n A^{n-1} \frac{dA}{dt}.
$$

Proof. We will prove this by induction. Let A be a matrix that commutes with its derivative. We will begin with the base case. Let $n = 1$

$$
\frac{d}{dt} A^1 = 1 A^0 \frac{d}{dt} A^1
$$

so we have

$$
\frac{d}{dt} A^n = n A^{n-1} \frac{dA}{dt}
$$

holds for $n = 1$.

Assume for some $n-1$ (where $n \in \{2, 3, \ldots\}$) that we have

$$\frac{d}{dt}A^{n-1} = (n-1)A^{n-2}\frac{dA}{dt}.$$

Now consider $\frac{d}{dt}A^n$

$$\begin{aligned}
\frac{d}{dt}A^n &= \frac{d}{dt}(A^{n-1}A) \\
&= (\frac{d}{dt}A^{n-1})A + A^{n-1}\frac{d}{dt}A \\
&= (n-1)A^{n-2}\frac{dA}{dt}A + A^{n-1}\frac{dA}{dt} \\
&= (n-1)A^{n-1}\frac{dA}{dt} + A^{n-1}\frac{dA}{dt} \\
&= nA^{n-1}\frac{dA}{dt}
\end{aligned}$$

So the assertion also holds true for n. Since we have that

$$\frac{d}{dt}A^n = nA^{n-1}\frac{dA}{dt}$$

is true for $n=1$ we see by induction that that holds for $n = 1, 2, \ldots$.

$\square$

Theorem 3.0.2. *Suppose an $n \times n$ matrix A commutes with its own integral,* $\left[A, \int_0^t A\ ds\right] = 0$*, then $x = e^{\int_0^t A\ ds}$ is a fundamental matrix solution to the system of differential equations $\frac{dx}{dt} = Ax$.*

Proof. Let A be a matrix that commutes with it's integral and let $x = e^{\int_0^t A\ ds}$. Consider $\frac{d}{dt}x$

$$\frac{d}{dt}(e^{\int_0^t A ds}) = \frac{d}{dt}\left(\sum_{k=0}^{\infty} \frac{(\int_0^t A \, ds)^k}{k!}\right)$$

$$= \sum_{k=0}^{\infty} \frac{d}{dt}\left(\frac{(\int_0^t A \, ds)^k}{k!}\right)$$

By Lemma 3.0.1 we have that

$$\frac{d}{dt}\left(\left(\int_0^t A \, ds\right)^k\right) = k\left(\int_0^t A \, ds\right)^{k-1} A$$

Putting this into the sum above we see that each term gets an extra factor of A.

$$\frac{d}{dt}x = \frac{d}{dt}(e^{\int_0^t A \, ds}) = \sum_{k=1}^{\infty} \frac{k\left(\int_0^t A \, ds\right)^{k-1} A}{k!}$$

$$= A\left(\sum_{k=0}^{\infty} \frac{\left(\int_0^t A \, ds\right)^k}{(k)!}\right)$$

$$= Ae^{\int_0^t A \, ds} = Ax$$

This shows us that x is a solution to the system of differential equations, because we have

$$\frac{d}{dt}x = Ax.$$

$\square$

If L_w commutes with its own integral, then by Theorem 3.0.2 this implies that $e^{\int_0^t L_w \ ds}$ is a fundamental matrix solution to our system of differential equations. A special case, and the case we will focus most of our attention on, is the case where all rate constants are equal. In this case the matrix we are studying is one where all the rate constants are constant numbers and equal or where all rate constants are set equal to a function that can be factored out of the matrix. As an aside; in the case where L_w it does not commute with it's integral, then we can approximate a general solution using the Peano-Baker series (2).

So we have access to the form of our general solution; $e^{\int_0^t L_w \ ds}$. However, it is not in a very useful form. We would like L_w to be diagonalized so we can have access to it's eigenvalues, as entries in the diagonal matrix D, and eigenvectors, as columns in the matrix B.

Definition 3.0.2. Exponential matrix

Let C be a $n \times n$ matrix with real or complex entries. The *exponential matrix* of C is given by

$$e^C = \sum_{k=0}^{\infty} \frac{1}{k!} C^k.$$

Definition 3.0.3. Diagonalizable Matrix

An $n \times n$ matrix A is said to be *diagonalizable* if there exists a matrix B that is non-singular and a diagonal matrix D such that $A = BDB^{-1}$.

Theorem 3.0.3. *Let C be a diagonalizable matrix. If $C = BDB^{-1}$ then*

$$e^C = Be^D B^{-1}.$$

Proof.

$$e^C = e^{BDB^{-1}}$$

$$= \sum_{n=0}^{\infty} \frac{BD^n B^{-1}}{n!}$$

$$= B \left(\sum_{n=0}^{\infty} \frac{D^n}{n!} \right) B^{-1}$$

$$= Be^D B^{-1}$$

So we arrive at our desired conclusion:

$$e^C = Be^D B^{-1}$$

$\square$

Lemma 3.0.4. *Let C be a diagonalizable matrix that commutes with it's integral with $C = BDB^{-1}$. If B is a constant matrix, then $x = Be^{\int_0^t D \ ds} B^{-1}$ is a fundamental matrix solution to the system of differential equations $\frac{dx}{dt} = Ax$.*

Proof. From Theorem 3.0.2 we have that $x = e^{\int_0^t C \ ds}$ is a solution. Starting from

there and doing some algebra we get

$$x = e^{\int_0^t C \ ds}$$

$$= e^{\int_0^t BDB^{-1} \ ds}$$

$$= e^{B \int_0^t D \ ds B^{-1}}$$

$$= Be^{\int_0^t D \ ds} B^{-1}$$

We use the fact that B is constant in the third equality and the final result comes from Theorem 3.0.3. $\square$

CHAPTER 4:

EIGENVALUES AND EXPLICIT SOLUTIONS

Now that we have established our system of equations and we know our general solution is of the form $e^{\int A} = Be^{\int D}B^{-1}$, we want have to describe B, B^{-1}, and D are. The first thing we have to do is find the eigenvalues, which will be the entries of D. To do this the graph in Figure 3.1 will need to be altered by treating the directed edges as an undirected edge by setting all the rate constants to be equal. This makes a path graph which can be seen in Figure 4.1. The results of this chapter can be found in reference (13).

Definition 4.0.1. Path Graph

A *path graph* or *linear graph* is a graph, for which the vertices can be listed in order such that the edges are $\{e_i, e_{i+1}\}$ with $i = 1, 2, ..., n-1$. All vertices have degree 2 except the two terminal vertices which have degree 1.
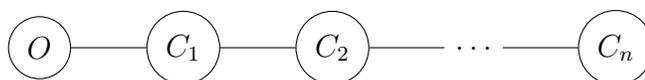
Figure 4.1 Taking 3.1 and forcing the rate constants to be equal establishes the directed edges as undirected edges. The vertices are still named to make clear to the reader that we will be intentionally still be using this to model the indiscernible n closed states of lysenin.

$$L_p = \begin{pmatrix} -k & k & & \dots & 0 \\ k & -2k & k & \dots & 0 \\ 0 & k & -2k & \dots & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & \dots & k & -2k & k \\ 0 & \dots & 0 & k & -k \end{pmatrix} = -k \begin{pmatrix} 1 & -1 & & \dots & 0 \\ -1 & 2 & -1 & \dots & 0 \\ 0 & -1 & 2 & \dots & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & \dots & -1 & 2 & -1 \\ 0 & \dots & 0 & -1 & 1 \end{pmatrix}$$

This is a Laplacian matrix of the path graph multiplied by $-k$. The Laplacian matrix of the path graph is well studied. So from here we will look into the Laplacian matrix of the path graph and establish the eigenvalues like we set out to do.

Definition 4.0.2. The *adjacency matrix* of a simple graph G is the $n \times n$ matrix $A = (A_{ij})$ where A_{ij} is equal to 1 when there is an edge in the graph G between vertex i and vertex j for $i \neq j$ and is equal to 0 when there is no edge.

Definition 4.0.3. The *degree matrix* D is a diagonal matrix with diagonal elements given by the number of edges incident to each vertex.

For the corresponding definition of directed multigraphs, see reference (13).

Definition 4.0.4. The *Laplacian matrix* L is defined as $L = D - A$.

In order to find the eigenvalues for the Laplacian of the path graph, we will take a step back and find eigenvalues for the cycle graph instead. First we need to establish some set up.

Definition 4.0.5. Union of Two Graphs

Let $G_1 \cup G_2$ be the *union of two graphs* G_1 and G_2, which have the same vertex set V and edge sets E_1 and E_2. This is the graph with vertex set V and edge set $E_1 \cup E_2$.

Definition 4.0.6. n-fold Multiplication of a Directed Multigraph

Let G be a directed multigraph with vertex set $\{v_1, ..., v_m\}$. Then G^n, for $n \geq 1$, is the directed graph with the same vertex set. Edges in G^n from v_i to v_j, $i, j \in \{1, ..., m\}$ are defined by oriented walks in G of length n (i.e. sequences of oriented edges $e_1 e_2 ... e_m$ where e_l is an edge from v_{i_l} to $v_{i_{l+1}}$ in G, $v_{i_l} = v_i$ and $v_{i_{m+1}} = v_j$). Note that G^0 is the graph with the identity matrix, I, as its adjacency matrix and thus has vertex set $\{v_1, .., v_m\}$ and a simple (directed) loop at each v_i.

Example 4.0.1. n-fold Multiplication of a Graph

Let G be the graph
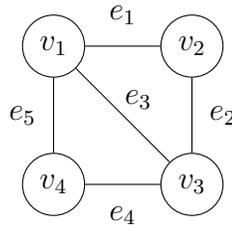
Then G^2 is the graph

If we have a graph G with an ordered set of vertices, we are able to find the adjacency matrix corresponding to that ordered set of vertices and eventually find the set of eigenvalues using the following statements.

Notation: Let $\mathbb{F}$ be an algebraically closed field and let $M_n(\mathbb{F})$ be the $\mathbb{F}$-algebra of $(n \times n)$ matrices with components in $\mathbb{F}$.
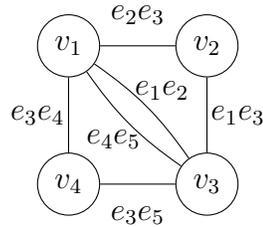
Proposition 4.0.1. Polynomial of a Graph

Let G be a graph and let $P(x)$ be a polynomial with coefficients $a_i \in \{0, 1, ...\}, i = 0, 1, ..., n$. Then there is a graph $P(G)$ whose adjacency matrix is $P(A)$. If we have a polynomial $P(x) = a_0 x^n + ... + a_n x^0$ then we can use $P(G) = a_0 G^n \cup ... \cup a_n G^0$. (13)

Theorem 4.0.1. *Let $A \in M_n(\mathbb{F})$. Let $\{\lambda_1, \lambda_2, ..., \lambda_n\} \subset \mathbb{F}$ be the set of eigenvalues for A and let $P(x)$ be a polynomial over the field $\mathbb{F}$. Then*

$$\det(P(A)) = \prod_{i=1}^{n} P(\lambda_i).$$

Proof. The fundamental theorem of algebra says that for the characteristic polynomial $\det(A - Ix) = (\lambda_1 - x)(\lambda_2 - x)...(\lambda_n - x) = \prod_{i=1}^{n}(\lambda_i - x)$ where each λ_i may be real or complex. Let $P(x) = c \prod_{l=1}^{m}(x - c_l)$ with $c, c_l \in \mathbb{F}$. Now we will use $P(A)$ as our new matrix and apply the determinant operation to it.

$$\det(P(A)) = \det\left(c \prod_{l=1}^{m}(A - c_l I)\right)$$

$$= c^n \prod_{l=1}^{m} \det(A - c_l I)$$

$$= c^n \prod_{l=1}^{m} \prod_{i=1}^{n}(\lambda_i - c_l)$$

$$= \prod_{i=1}^{n}\left(c \prod_{l=1}^{m}(\lambda_i - c_l)\right)$$

$$= \prod_{i=1}^{n}(P(\lambda_i))$$

$\square$

In Theorem 4.0.1 we were able to solve for what the determinant for $P(A)$ was. However, we are most interested in the eigenvalues, so in the following theorem we will solve for the eigenvalues.

Theorem 4.0.2. *Let $A \in M_n(\mathbb{F})$ and $P \in \mathbb{F}[x]$. If $\lambda_1, \lambda_2, ..., \lambda_n$ are the eigenvalues of A, then $P(\lambda_1), P(\lambda_2), ..., P(\lambda_n)$ are the eigenvalues of $P(A)$.*

Proof. Let γ be an indeterminate over $\mathbb{F}$. Define $\mathbb{K}$ as the algebraic closure of $\mathbb{F}[\gamma]$. Let $H \in \mathbb{K}[x]$ with definition $H(x) = \gamma - P(x)$. Consider $H(A)$

$$H(A) = \gamma I - P(A)$$

and so

$$\det(H(A)) = \det(\gamma I - P(A)) \tag{4.1}$$

Note that $\det(\gamma I - P(A))$ is the characteristic polynomial of $P(A)$, which is what defines the eigenvalues of $P(A)$. Since $\mathbb{K}$ is an algebraically closed field and we have $A \in M_n(\mathbb{F}) \subset M_n(\mathbb{K})$ and $H \in \mathbb{K}[x]$ we can use Theorem 4.0.1 to get

$$\det(H(A)) = \prod_{i=1}^{n} H(\lambda_i). \tag{4.2}$$

Combining Equations 4.1 and 4.2 and then apply H we get

$$\det(\gamma I - P(A)) = \prod_{i=1}^{n} H(\lambda_i)$$
$$= \prod_{i=1}^{n} (\gamma - P(\lambda_i))$$

So the characteristic polynomial of $P(A)$ is given by $S(\gamma) = \prod_{i=1}^{n}(\gamma - P(\lambda_i))$, whose roots are $P(\lambda_1), P(\lambda_2), ..., P(\lambda_n)$. Thus $P(\lambda_1), P(\lambda_2), ..., P(\lambda_n)$ are the eigenvalues of $P(A)$. $\qquad\square$

Example 4.0.2. n-fold Multiplication of a Graph

Let $\vec{C}_n$ be the cycle graph below.
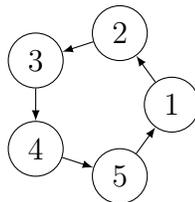
The above image is a circuit of five vertices which directed edges. By raising this graph to powers we make new connections. To be specific, if we multiply the graph to itself, we get the following. Notice how the vertices are now connecting the vertex that was two spaces away.

$$\vec{C}_5 \cdot \vec{C}_5 = \quad \bullet \quad = \quad = \vec{C}_5^2$$
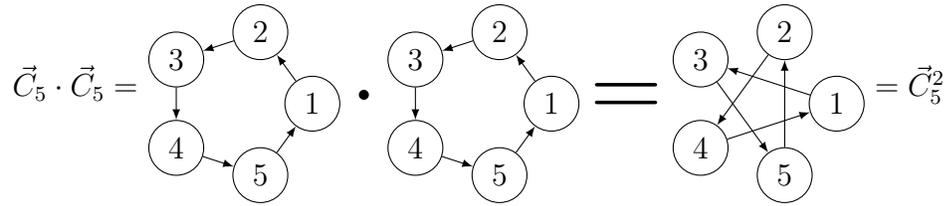
I will continue until I arrive at $\vec{C}_5^4$.

$$\vec{C}_5^2 \cdot \vec{C}_5 = \quad \bullet \quad = \quad = \vec{C}_5^3$$

$$\vec{C}_5^3 \cdot \vec{C}_5 = \quad \bullet \quad = \quad = \vec{C}_5^4$$
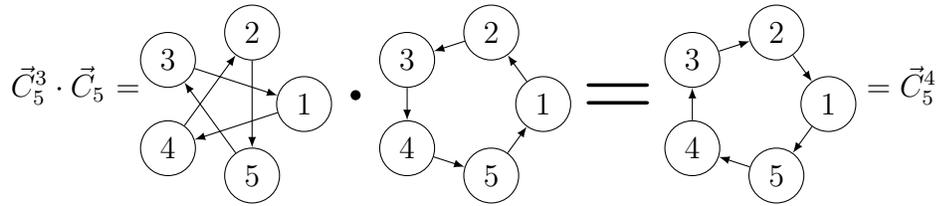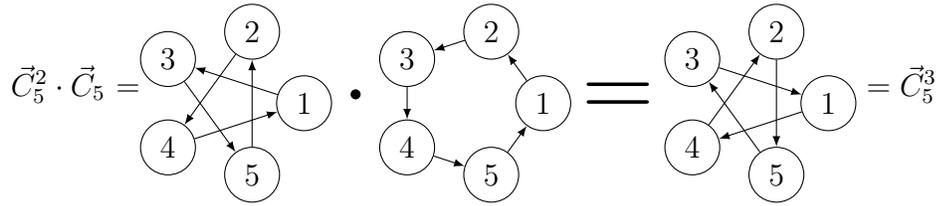
If we do this for general n we see that $\vec{C}_n^{n-1}$ is the graph of $\vec{C}_n$ with edge orientations reversed. Using what we have established above, what we really want is the undirected cycle graph C_n, but to get it we will set up $C_n = \vec{C}_n \cup \vec{C}_n^{n-1}$ where $\vec{C}_n^{n-1}$ is a directed cycle graph which has been specially set up to have edges that are complimentary and oppositely oriented to $\vec{C}_n$.
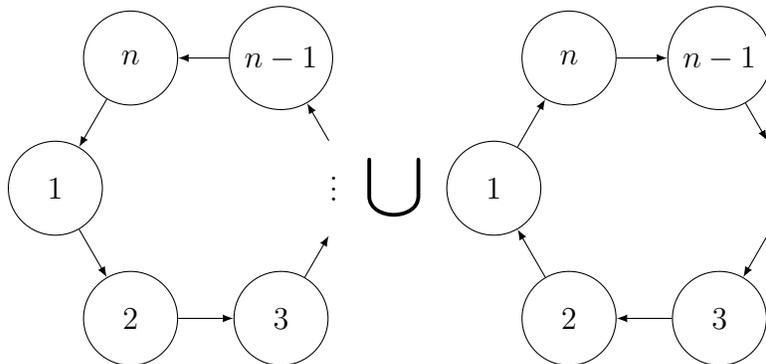
The adjacency matrix of $\vec{C}_n$ is the following.

$$A = \begin{pmatrix} 0 & 1 & 0 & \ldots & 0 \\ 0 & 0 & 1 & \ldots & 0 \\ 0 & 0 & 0 & \ldots & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & \ldots & & 0 & 1 \\ 1 & \ldots & 0 & 0 & 0 \end{pmatrix}$$

We want to compute powers of this adjacency matrix, we will begin with an example.

Example 4.0.3. Raising the Adjacency Matrix to Powers

Before moving into a proof that will allow us to assume the general case of what raising A to a power does, I will offer a small example first to motivate the proof.

$$A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}$$

The matrix above is an adjacency matrix for a directed cycle with three vertices. Notice that it is also a permutation matrix and because of this, when we raise A to the third power, we get the identity matrix I. When we raise it to the fourth power, we get back A.

$$A^2 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

Now we will cube A by multiplying our new matrix by A.

$$A^3 = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Notice how the rows permute themselves downward with the final row being thrown to the top of the matrix. I will now move into proving that this will be the case in general so we can use this permutation of rows to our advantage.

Lemma 4.0.3. *Let*

$$A = \begin{pmatrix} 0 & I_{n-1} \\ I_1 & 0 \end{pmatrix}$$

and e_i be a canonical basis vector in $\mathbb{R}^n$ with $i = 1, 2, \ldots, n$ then

$$e_i A = e_{i+1}$$

for $i = 1, ..., n-1$ and

$$e_n A = e_1$$

Proof. Let i be an integer between 1 and $n-1$ and e_i be a general basis vector. Let A be represented by the row vector of basis vectors $(e_n, e_1, e_2...e_{n-1})$.

Multiplying these two things together we get

$$e_i A = e_i(e_n, e_1, e_2...e_{n-1}) = (e_i e_n, e_i e_1, ..., e_i e_i, ..., e_i e_{n-1}) = (0, 0, ..., 1, ..., 0)$$

We also know that the exact position of our 1 is the $i+1$ spot and we know this because e_i is in the $i+1$ position in the vector we multiplied e_i by, it will shift e_i over by one index each time we do the multiplication. So $e_i A = e_{i+1}$ for $i = 1, 2, ..., n-1$.

The final thing we must address is the edge case of e_n.

$$e_n A = e_n(e_n, e_1, e_2...e_{n-1}) = (e_n e_n, e_n e_1, ..., e_n e_{n-1}) = (1, 0, ..., 0) = e_1$$

$\square$

Theorem 4.0.4. *Let*

$$A = \begin{pmatrix} 0 & I_{n-1} \\ I_1 & 0 \end{pmatrix}$$

then

$$A^k = \begin{pmatrix} 0 & I_{n-k} \\ I_k & 0 \end{pmatrix}$$

for $k = 1, 2, \ldots, n-1$

Proof. We will proceed by induction. First we need to tackle the base case and in this situation the base case will be A^k where $k = 0$ which is the identity matrix and that is what A^k predicts so the base case has been shown to work.

We assume the following equality holds for some j.

$$A^j = \begin{pmatrix} 0 & I_{n-j} \\ I_j & 0 \end{pmatrix}$$

Now we will show this is true for $j + 1$.

$$A^{j+1} = A^j A = \begin{pmatrix} 0 & I_{n-j} \\ I_j & 0 \end{pmatrix} \begin{pmatrix} 0 & I_{n-1} \\ I_1 & 0 \end{pmatrix}$$

Note that we can write A^j and A in terms of e_i:

$$A^j A = \begin{pmatrix} e_{n-j} \\ \vdots \\ e_n \\ e_1 \\ \vdots \\ e_j \end{pmatrix} (e_n, e_1, e_2, ..., e_{n-1})$$

By Lemma 4.0.3 we have $e_i A = e_{i+1}$ and $e_n A = e_1$, so we get

$$A^j A = \begin{pmatrix} e_{n-j} \\ \vdots \\ e_n \\ e_1 \\ \vdots \\ e_j \end{pmatrix} (e_n, e_1, e_2, ..., e_{n-1}) = \begin{pmatrix} e_{n-(j+1)} \\ \vdots \\ e_n \\ e_1 \\ \vdots \\ e_{j+1} \end{pmatrix}$$

$\square$

Theorem 4.0.5. *Eigenvalues of C_n*

Let A be the adjacency matrix of the undirected cycle graph C_n. The eigenvalues, λ_j where $0 = \lambda_1 < \lambda_2 < \ldots < \lambda_n$ are given by

$$\lambda_j = \eta_j + \eta_j^{-1}$$

where $\eta_j = e^{\frac{2\pi ij}{n}}$, for $j = 0, 1, \ldots, n-1$. Note that $\eta_j^n = 1$.

Proof. Let P be a polynomial defined as

$$P(x) = x + x^{n-1}$$

.

Let $\vec{C}_n$ be the directed cycle graph with n vertices and let A be it's adjacency matrix. If we apply P to A we get

$$P(A) = A + A^{n-1}$$

$$= \begin{pmatrix} 0 & I_{n-1} \\ I_1 & 0 \end{pmatrix} + \begin{pmatrix} 0 & I_1 \\ I_{n-1} & 0 \end{pmatrix}$$

$$= \begin{pmatrix} 0 & 1 & 0 & \ldots & 1 \\ 1 & 0 & 1 & & 0 \\ \vdots & & & \ddots & \vdots \\ 1 & 0 & \ldots & 1 & 0 \end{pmatrix}$$

$P(A)$ is the adjacency matrix for the undirected cycle graph, C_n, so $P(\vec{C}_n) = C_n$. It is well known that the eigenvalues of $\vec{C}_n$ is given by

$$\eta_j = e^{2\pi i j / n}.$$

For further reading and proof of these eigenvalues, see reference (4). Note that $\eta_j^n = 1$ for $j = 0, 1, \ldots n - 1$. By Theorem 4.0.2 we have that $P(\vec{C}_n) = C_n$ has the eigenvalues $P(\eta_j)$, which gives us

$$P(\eta_j) = \eta_j + \eta_j^{n-1}$$
$$= \eta_j + \eta_j^n \eta_j^{-1}$$
$$= \eta_j + \eta_j^{-1}$$

So the eigenvalues for the undirected cycle graph are (4)

$$\lambda_j = \eta_j + \eta_j^{-1},$$

for $j = 0, 1, \ldots, n - 1$

$\square$

Theorem 4.0.6. *Let L be the Laplacian matrix of undirected cycle graph, C_n. The eigenvalues, λ_j where $0 = \lambda_1 < \lambda_2 < \ldots < \lambda_n$, and their corresponding eigenvectors, v_j, are given by*

$$\lambda_j = 2 - \eta_j - \eta_j^{-1}$$

and

$$v_j = (1, \eta_j, \eta_j^2, ..., \eta_j^{n-1})^T$$

where $\eta_j = e^{\frac{2\pi i j}{n}}$, for $j = 0, 1, \ldots, n-1$. Note that $\eta_j^n = 1$.

Proof. Let L be the Laplacian matrix of the undirected cycle graph, C_n with adjacency matrix A and degree matrix D. In Theorem 4.0.5 we showed that the eigenvalues of A are

$$\gamma_j = \eta_j + \eta_j^{-1},$$

where $\eta_j = e^{\frac{2\pi i j}{n}}$ for $j = 0, 1, \ldots, n-1$. The corresponding eigenvectors are

$$v_j = (1, \eta_j, \eta_j^2, ..., \eta_j^{n-1})^T,$$

which can be found in *Spectra of Graphs* (4).

Consider Lv_j,

$$Lv_j = (D - A)v_j$$
$$= (2I - A)v_j$$
$$= 2Iv_j - Av_j$$
$$= 2v_j - \gamma_j v_j$$
$$= (2 - \gamma_j)v_j$$
$$= \lambda_j v_j$$

So we conclude that the eigenvalues of the L are $\lambda_j = 2 - \gamma_j = 2 - \eta_j - \eta_j^{-1}$ with the corresponding eigenvectors v_j. $\qquad\square$

Theorem 4.0.7. *Let L be the Laplacian matrix of the path graph P_n. The eigenvalues, λ_j where $0 = \lambda_1 < \lambda_2 < ... < \lambda_n$, and their corresponding eigenvectors, v_j, are given by*

$$\lambda_j = 2 - \eta_j - \eta_j^{-1}$$

and

$$v_j = (1 + \eta_j^{2n-1}, ..., \eta_j^k + \eta_j^{2n-1-k}, ..., \eta_j^{n-1} + \eta_j^n)^T$$

where $\eta_j = e^{\frac{\pi i j}{n}}$, for $j = 0, 1, \ldots, n-1$. Note that $\eta_j^{2n} = 1$.

Proof. The entries of L are

$$
a_{l,k} =
\begin{cases}
1 & \text{if } k = l \text{ and } l = 1 \text{ or } l = n \\
2 & \text{if } k = l, \\
-1 & \text{if } k = l - 1 \text{ or } k = l + 1, \\
0 & \text{otherwise}
\end{cases}
$$

When we multiply L and v_j the l-th entry of the resulting vector is

$$
(Lv_j)(l) = \sum_{k=1}^{n} a_{l,k} v_j(k)
$$

For the middle rows of L $(0 < l < n)$

$$
\begin{aligned}
(Lv_j)(l) &= a_{l,l-1} v_j(l-1) + a_{l,l} v_j(l) + a_{l,l+1} v_j(l+1) \\
&= -(\eta_j^{l-1} + \eta_j^{2n-1-(l-1)}) + 2(\eta_j^{l} + \eta_j^{2n-1-l}) - (\eta_j^{l+1} + \eta_j^{2n-1-(l+1)}) \\
&= (2 - \eta_j - \eta_j^{-1})(\eta_j^{l} + \eta_j^{2n-l-1}) \\
&= \lambda_j v_j(l)
\end{aligned}
$$

For the top row of L we have $(l = 1)$

$$
\begin{aligned}
(Lv_j)(1) &= a_{1,1} v_j(1) + a_{1,2} v_j(2) \\
&= (1 + \eta_j^{2n-1}) - (\eta_j + \eta_j^{2n-2}) \\
&= (2 - \eta_j - \eta_j^{-1})(1 + \eta_j^{2n-1}) \\
&= \lambda_j v_j(1)
\end{aligned}
$$

For the top row of L we have $(l = n)$

$$
\begin{aligned}
(Lv_j)(n) &= a_{n,n-1}v_j(n-1) + a_{n,n}v_j(n) \\
&= -(\eta_j^{n-2} + \eta_j^{2n-1-(n-2)}) + (\eta_j^{n-1} + \eta_j^n) \\
&= (2 - \eta_j - \eta_j^{-1})(\eta_j^{n-1} + \eta_j^n) \\
&= \lambda_j v_j(n)
\end{aligned}
$$

$\square$
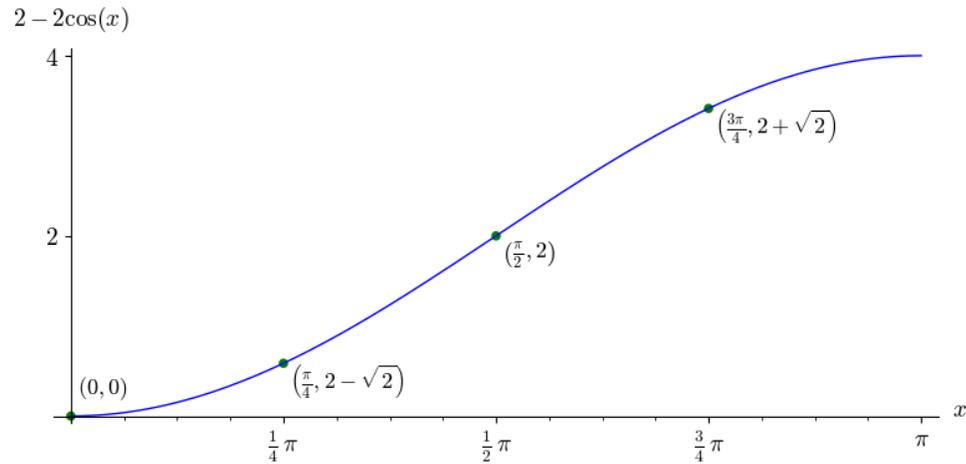
Figure 4.2 **The points on the plot are the eigenvalues are for a 4×4 Laplacian matrix of the path graph.**

We have found the eigenvalues and eigenvectors for L_p, so we can perform eigen-decomposition for L_p. Meaning we can write

$$L_p = BDB^{-1}$$

where B is the matrix with eigenvectors as it's columns and D is the diagonal matrix with eigenvalues. Since L_p is diagonalizable and B is a constant matrix then by Lemma 3.0.4

$$P(t) = B e^{\int_0^t -kD \ ds} B^{-1} P^0$$

is a solution to the differential system of equations $\frac{dP}{dt} = L_p P$, where $P(0) = p^0$. You may notice that we are missing one vital piece, namely B^{-1}. We could not determine a general formula for the inverse matrix of B. However, it is possible to manually compute (with aid from a computer) the inverse of B for a reasonably small n, which

we will do when utilizing this solution.

CHAPTER 5:

RESULTS AND DATA
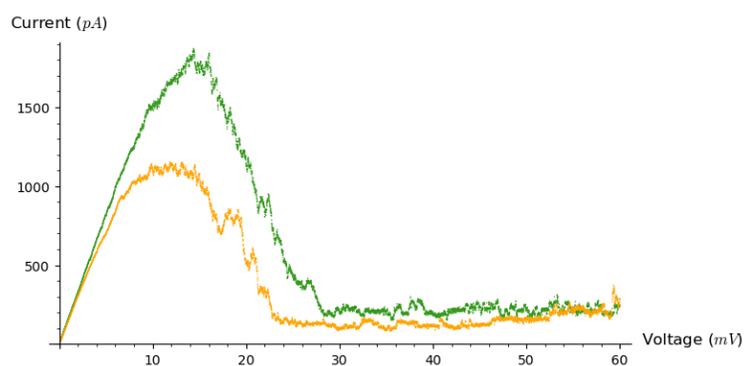
Figure 5.1   A hysteresis graph for a population of lysenin channels. The green curve shows how ionic current decreases as voltage is increased indicating that the channels are gating. The yellow curve indicates how the current increases as the voltage decreases indicating that the channels are in a conducting state.

In this section of the thesis I will be describing the practical analysis of data using the method described in the previous chapters. Note the data shown in Figure 5.1 was used to make the plot shown in Figure 5.2. The hysteresis plot is a very typical one that displays a slow change in voltage and the ionic currents changing in response. At each change in voltage the channels are allowed to relax and reach full equilibrium. This is called a 'static equilibrium' and is an uncommon feature in molecular biophysics; see Equation 5.2.
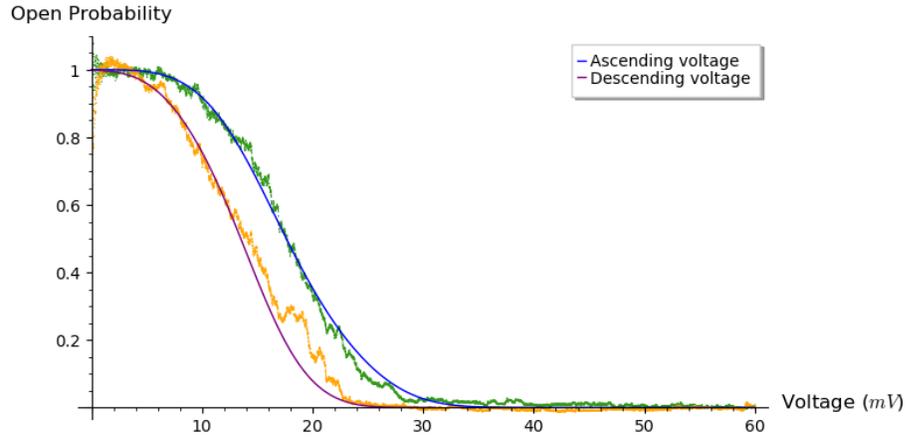
Figure 5.2 **A typical experiment's open probability data being modeled by the strategy researched in this thesis. The experiment can be broken into two parts. The first part, ascending voltage, is where the voltage the channels are exposed to increases slowly so they reach equilibrium at each new potential so that eventually all channels will be gated. The second part, descending voltage, is where the channels are slowly re-opened by decreasing the voltage very slowly.**

$$P_{Open}(t) = \frac{k_+ e^{-t(k_+ + k_-)}}{k_+ + k_-} + \frac{k_-}{k_+ + k_-} \tag{5.1}$$

$$\lim_{t \to \infty} P_{Open}(t) = \frac{k_-}{k_+ + k_-} \tag{5.2}$$

To model the data as seen in Figure 5.2 we begin with the data that makes up a hysteresis; the voltage and current. With this raw data we compute the open probability by taking I_m (Equation 5.3) and dividing it by the theoretical maximum current (Equation 5.5). Both of these things count the the channels; I_m signifies the number of open channels at a given voltage and I_{max} is the total population of channels.

$$I_m = O \cdot G_0 \cdot V \tag{5.3}$$

$$I_0 = G_0 \cdot V \tag{5.4}$$

$$I_{max} = I_0 \cdot N \tag{5.5}$$

The next step is to decide on the number of closed states we want to use to model the data; in most of the modeling I did for this thesis I assumed there were between 2 and 4 closed states. We also need to decide upfront what the rate constant will be. For the data analyzed for this thesis simple, often single term, polynomials were chosen. For example, V or t, but more terms are not forbidden. There is even the possibility of using logarithmic, exponential, or other function types.

In Figure 5.2 we assumed that there were 3 closed states for lysenin and took the rate constant to be x^3 for the ascending voltage and the rate constant to be x^4 for the descending voltage. With those parameters we generated our eigenvalues and eigenvectors and computed the inverse of the matrix of eigenvectors to find the solution to our differential equations. We have two different initial conditions for ascending and descending voltages. For the ascending voltage we assume that every channel is open, so the vector is $P^0 = (1, 0, 0, 0)^T$. For the descending voltage we assume that there is an equal distribution among the closed states, so the vector is $P^0 = (0, \frac{1}{3}, \frac{1}{3}, \frac{1}{3})^T$. With these initial conditions we can get our unique solution for open probability, $P_{Open}(t)$.

There is a slight problem with this function in that we needed to shift and scale the resulting function to fit in the range of 0 and 1. If we approach a continuous model by

increasing the number of states we would naturally get this shifting and stretching for ascending voltage. Alternatively if we had a different model that allowed for the rate constants to be different from one another as voltage changes (i.e. allowing the rate constants to reflect changes in the energy landscape). Having one constant means that we have lost some resolution to how the model fits the data.

The probability functions graphed in Figure 5.2 are modified in two ways. As previously mentioned they have to be stretched and the descending voltage graph needs to be reversed (the plot is in terms of voltage and not in time, so we have to adjust our plot accordingly). In Figure 5.3 we have plotted the versions of the open probability for ascending and descending voltage that have not been scaled or shifted (the descending voltage probability has been flipped to account for the voltage plot). You can see that the probability functions have an asymptote at $\frac{1}{4}$ and more generally $\frac{1}{n}$ where we have $n-1$ closed states. This defect in the model is expected, because the rate constants are equal and so the probability of being in any state will be equally likely as time goes on.
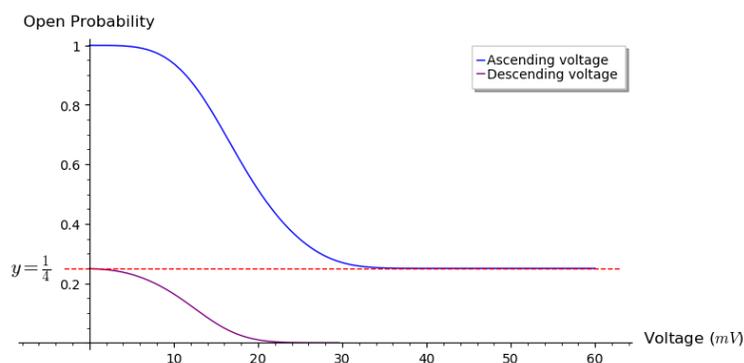
Figure 5.3   Probability functions plotted in terms of voltage with no scal-ing or shifting

The specific modifications we made to the open probability functions to generate

the graph in Figure 5.2 are

- **Ascending voltage**

$$\frac{nP_{Open} - 1}{n - 1}$$

- **Descending voltage**

$$nP_{Open}$$

Physically we understand that the channels will go all the way to the last possible closed position as voltage is increased and when we allow the channels reach equilibrium. Figure 5.4 displays how different a fit can be just by altering the initial condition for the descending voltage from $P^0 = (0, \frac{1}{3}, \frac{1}{3}, \frac{1}{3})^T$ to $P^0 = (0, 0, 0, 1)^T$. I would like to bring your attention specifically to the fact that the area enclosed in the curves in Figure 5.4 is larger than in Figure 5.2 which indicates that the channels do take more time to open when they are able to fully reach equilibrium in the final closed position.
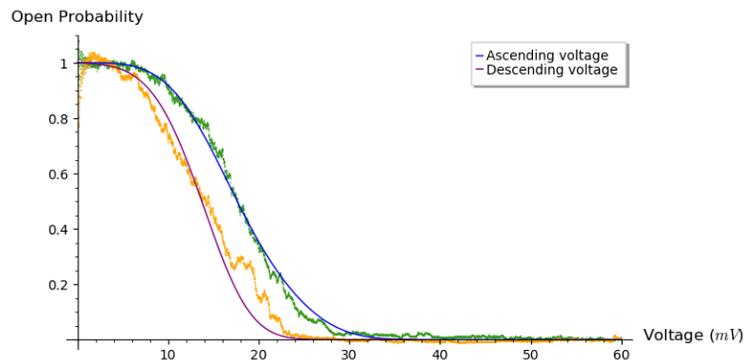
Figure 5.4 Open probability voltage plot where the descending voltage initial condition is taken to be $(0, 0, 0, 1)^T$ and the $k = t^3$. No changes to ascending voltage shown in Figure 5.2

CHAPTER 6:

FURTHER THOUGHTS

Naturally when doing research more and more areas to investigate show up. There are handful that came to our attention the course of researching this thesis that we will mention here which would be natural next steps to further this line of research that we could not get to.

First, in our model that we came to at the end of Chapter 4 there is a single-variable function k. Determining a 'good' function that models the data appropriately is primarily guess-and-check. There are a lot of potential functions that one could try. We limited our choices to monomial terms $(k = t^j)$, but other functions could be considered like $k = e^t$ or $k = t^5 + t^3 + 4$. Determining a better way to choose a function is something worth looking into. Also if we could determine an optimal function type (polynomial, exponential or something else) that would work best that would also be of interest.

Next, we were able to diagonalize the matrix L_p at the end of Chapter 4, but we were not able to explicitly write what the inverse of B would be. We have reason to believe that there is a general form for it's inverse due to the Vandermonde nature of the matrix of eigenvectors B. If we had B^{-1} we would be able to write the closed form solution for n-closed states.

Lastly, in Chapter 5 we had to shift and scale the resulting probability functions

to fit the data. The cause of that manipulation of the graph was due to their only be one rate constant for the Markov process, which led to each state being equally likely as t increases. This naturally would prompt us to look at increasing the number of rate constants in our model. This would be a completely different model, but even increasing the number of rate constants to two is thought to remove the need for this shifting and scaling. It would also be an interesting topic to investigate as it would provide an even more general model.

CHAPTER 7:

CONCLUSION

Understanding the physical role of molecules is essential, not only for making connections to known biological functionalities, but also for deciphering novel and yet uncovered features. Understanding how memory may occur at a molecular level is the ultimate goal of studying molecules which may exhibit memory through hysteretic behavior. Lysenin is one of the best molecules for studying hysteretic behavior in cellular biology due to the fact that it is easy to work with in an experimental setting, relatively inexpensive, and it displays that it remembers its past history for many hours during experimentation. Because of these reasons, and others, we have chosen to invest in improving the modeling used to understand it within thesis.

Typically when modeling ion channels, which lysenin has been compared to in the biophysical literature, the model only recognizes conducting and non-conducting states, see Figure 2.4. However, it is thought that lysenin may have more electrically indiscernible states due to the fact that it does display a strong hysteresis in conductance when exposed to oscillatory voltage stimuli.

The approach described in this thesis reasonably models the hysteretic behavior of lysenin by applying well-studied mathematical concepts in novel ways. We were able to take the two-state system and modify it to be more general with n states, closed or open, since both conducting and non-conducting states are electrically indiscernible.

The model discussed in this paper also does well to mathematically explain, in part, the energy-landscape as it changes with the pores. As the channels are exposed to external stimuli (such as voltage) their preferred state changes and they alter their confirmation. This is reflected in the multiplicity of the eigenvalues. Since the eigenvalues are never repeated, see Figure 4.2, this models how the energetics of the system change.

Studying this mechanism is important because the hysteresis is linked to memory. Since memory is typically thought of as a product of a more complex system, like a population of cells or even a brain, it is important to study a molecule that may exhibit memory. In addition, lysenin behaves similarly to ion channels in our brain, muscles, and all cells, and scientists may look for similar hysteretic behavior in the functionality of those transporters and attempt to identify its physiological relevance. Given that lysenin channels are electrically controlled, future explorations may focus on its use it as memory in devices that currently use solid-state elements, or even as an interface between neuron-based systems and artificial systems.

REFERENCES

[1] A, T., GT, L., AG, E., AND A., M. Computing rates of markov models of voltage-gated ion channels by inverting partial differential equations governing the probability density functions of the conducting and non-conducting states. *Mathematical Biosciences 277*, 126–135.

[2] BAAKE, M., AND SCHLAEGEL, U. The peano-baker series, 2012.

[3] BENNEKOU, P., JENSEN, L. R., BARKSMANN, T. L., KRISTENSEN, B. I., AND CHRISTOPHERSEN, P. Voltage activation and hysteresis of the non-selective voltage-dependent channel in the intact human red cell. *Bioelectrochemistry 62*, 2, 181–185.

[4] BROUWER, A. E., AND HAEMERS, W. H. *Spectra of Graphs*. Springer, New York, NY, 2012.

[5] BRYANT, S., CLARK, T., THOMAS, C., WARE, K., BOGARD, A., CALZA-CORTA, C., PRATHER, D., AND FOLOGEA, D. Insights into the voltage regulation mechanism of the pore-forming toxin lysenin. *Toxins 10*, 8, 334.

[6] COMMONS, W. File:eisenia fetida hc1.jpg — wikimedia commons, the free media repository, 2016.

[7] COMMONS, W. File:lysenin action mechanism.png — wikimedia commons, the free media repository, 2018.

[8] FOLOGEA, D., KRUEGER, E., MAZUR, Y., STITH, C., OKUYAMA, Y., HENRY, R., AND SALAMO, G. Bi-stability, hysteresis, and memory of voltage-gated lysenin channels. *Biochim Biophys Acta. 12*, 2933–2939.

[9] FOLOGEA, D., SHRESTHA, N., THOMAS, C., RICHTSMEIER, D., PU, X., TINKER, J., AND BRYANT, S. L. Stochastic sensing of angiotensin ii with lysenin channels. *Scientific Reports 7*, 2448.

[10] IDE, T., AOKI, T., TAKEUCHI, Y., AND YANAGIDA, T. Lysenin forms a voltage-dependent channel in artificial lipid bilayer membranes. *Biochem Biophys Res Commun. 346*, 288–292.

[11] LEE, A. G. A paddle in oil. with permission, license number 4932021357519.

[12] MILLHAUSER, G. L., SALPETER, E. E., AND OSWALD, R. E. Diffusion models of ion-channel gating and the origin of power-law distributions from single-channel recording. *Proceedings of the National Academy of Sciences of the United States of America 85*, 5, 1503–1507.

[13] SACHS, H., DOOB, M., AND CVETKOVIC, D. M. *Spectra of graphs: Theory and application.* New York: Academic Press, 1980.