A DOUBLY STOCHASTIC POISSON PROCESS MODEL FOR WAKE-SLEEP CYCLING

DISSERTATION

Presented in Partial Fulfillment of the Requirements for the Degree Doctor of Philosophy in the Graduate School of the Ohio State University

By
Badal Joshi, B.Sc., M.S.(Physics), M.S.(Mathematics)
Graduate Program in Mathematics

The Ohio State University
2009

Dissertation Committee:
Professor Janet Best, Advisor
Professor Avner Friedman
Professor Saleh Tanveer
ABSTRACT

In adult mammalian species, sleep is composed of several micro-sleep bouts which are exponentially distributed and micro-wake bouts which follow a power law. In infant rats both sleep and wake bouts follow an exponential law, only as the animal develops the wake bout times develop the heavy tail of a power law.

Based on a survey of experimental findings, we identify the populations of neurons responsible for the state change behavior and propose a connectivity diagram consistent with the known neurophysiology. We suggest a new general modeling approach where individual populations of neurons are modeled as Poisson processes whose rates are stochastic processes and satisfy a system of stochastic differential equations. We also suggest a canonical map from the connectivity diagram to the Poisson process model.

The analysis of the stochastic dynamical system is based on an appropriate deterministic approximation. The deterministic dynamical system is analyzed using standard results and these results are then used to make predictions about the stochastic system. The model shows the appropriate behavior of random switching between sleep and wake and how the probability of this switching behavior changes with age.
We find that a necessary condition for the bout distribution in the stochastic system to change from exponential to power law is that the deterministic system has a bifurcation from one stable fixed point to two stable fixed points.

Finally, in order to compare the theoretical predictions with experimental data we develop algorithms for parameter estimation and for comparison of the simulated and the experimental data.
To

My parents

With Love
ACKNOWLEDGMENTS

I would like to thank my advisor, Dr. Janet Best who suggested the thesis problem and who has always been a source of support and encouragement. Because of her guidance, I was introduced to many new areas and prominent researchers in mathematical biology. I also want to thank the dissertation committee members Dr. Avner Friedman and Dr. Saleh Tanveer. I first came in contact with Dr. Tanveer when I was changing my field of study from physics to mathematics. I am thankful for his guidance and encouragement during this process and later during my time as a graduate student. I am thankful to Dr. Boris Pittel and Dr. Bjorn Sandstede for positively influencing my mathematical education. I want to especially mention Dr. Gheorghe Craciun, Dr. Laura Kubatko, Dr. Dennis Pearl, Dr. Ralf Bundschuh and Dr. Anastasios Matzavinos among others who have all helped me in my education and research.

Besides all the love and encouragement, I want to thank my parents for always supporting my decisions and fostering my early interest in science. Thanks to my parents and my sister Ruchi for always being present despite the distance. I have met many exceptional people at Ohio State, in particular I am grateful to Yogesh Srivastava for being a close friend and for those endless discussions on science and everything else. I also thank my good friends Martin Nikolov and Lore Thaler.
VITA

2004-Present .......................... Graduate Teaching Assistant, Mathematics Department,
The Ohio State University

2000-2004 ............................. Graduate Teaching Assistant, Physics Department
The Ohio State University

2008 .................................. M.S. in Mathematics,
The Ohio State University

2004 .................................. M.S. in Physics,
The Ohio State University

1996 ................................. M.B.A., Gujarat University, India

1994 ................................. B.Sc. in Physics, Gujarat University, India
PUBLICATIONS

Research Papers


# TABLE OF CONTENTS

Abstract ......................................................... ii
Dedication ......................................................... iv
Acknowledgments .................................................. v
Vita ................................................................. vi
List of Figures ...................................................... xi
List of Tables ...................................................... xiv

<table>
<thead>
<tr>
<th>CHAPTER</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Introduction ............................................... 1</td>
</tr>
<tr>
<td>2</td>
<td>Biological background and description of the basic problem .. 3</td>
</tr>
<tr>
<td>3</td>
<td>Model ........................................................... 7</td>
</tr>
<tr>
<td></td>
<td>3.1 Introduction ............................................... 7</td>
</tr>
<tr>
<td></td>
<td>3.2 The main idea ............................................... 10</td>
</tr>
<tr>
<td></td>
<td>3.3 Definitions and Preliminaries ............................ 10</td>
</tr>
<tr>
<td></td>
<td>3.4 SDE system .................................................. 14</td>
</tr>
<tr>
<td></td>
<td>3.5 Interpretation of the SDE system ....................... 16</td>
</tr>
<tr>
<td></td>
<td>3.6 Graphical representation ................................. 17</td>
</tr>
<tr>
<td></td>
<td>3.7 Explicit forms of inhibition and excitation functions ... 18</td>
</tr>
<tr>
<td></td>
<td>3.8 Circuit for sleep-wake switch ............................. 23</td>
</tr>
<tr>
<td>4</td>
<td>Deterministic system ......................................... 28</td>
</tr>
<tr>
<td></td>
<td>4.1 Derivation of the deterministic system .................. 28</td>
</tr>
<tr>
<td></td>
<td>4.2 Application of Results to the Sleep-Wake Model ......... 35</td>
</tr>
</tbody>
</table>
5 Ruling out closed orbits and chaotic behavior .................................. 37
6 Analysis of the Two Component Deterministic System ................. 39
   6.1 Reduced system with zero self-excitation .......................... 40
   6.2 General case of non-negative self-excitation ....................... 43
   6.2.1 Estimates for number of steady state solutions ................. 46
   6.2.2 Wake-sleep classification for steady state solutions .......... 52
   6.3 Bifurcation analysis of steady state solutions for two com-
   ponent model ........................................................................ 56
   6.3.1 Bifurcation parameter is self-excitation $\alpha = \alpha_{11} = \alpha_{22}$ ..... 56
   6.3.2 Bifurcation parameter is mutual inhibition $\beta = \beta_{12} = \beta_{21}$ ... 59
   6.3.3 Bifurcation diagrams for the case when $\alpha_{11} \neq \alpha_{22}$ ........ 60
7 Simulation results of the stochastic model for the two component system 66
   7.1 Zero Self-excitation .......................................................... 66
   7.2 Positive Self-excitation ....................................................... 68
8 Analysis of the Three Component Deterministic System ............... 74
   8.1 Two component mutually excitatory system ....................... 74
   8.2 Effect of external excitation on two component wake-sleep
   system .................................................................................. 76
   8.2.1 Zero self-excitation ......................................................... 77
   8.2.2 Positive self-excitation ................................................... 80
   8.3 Bifurcation parameter is mutual excitation parameter - Ef-
   fect of development on wake distribution ................................ 82
9 Simulation results of the stochastic model for the three component system 88
   9.1 Firing rate plots ................................................................. 88
   9.2 Survivor plots ..................................................................... 91
10 Mechanism for Power Law .......................................................... 95
   10.1 Biased random walk model ............................................... 95
   10.2 Connection with the neurophysiological model ................... 96
   10.3 Solution of the biased random walk ................................... 97
## LIST OF FIGURES

<table>
<thead>
<tr>
<th>FIGURE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1</td>
<td>Graphical representation of a mixed excitatory-inhibitory system</td>
</tr>
<tr>
<td>3.2</td>
<td>Commonly occurring motifs of mutual excitation and mutual inhibition</td>
</tr>
<tr>
<td>3.3</td>
<td>The full circuit diagram for sleep-wake switch</td>
</tr>
<tr>
<td>6.1</td>
<td>Nullclines in phase plane ( (\lambda_1, \lambda_2) ) for mutual inhibition parameter ( \beta = 0.5 ). The values of self-excitation parameter ( \alpha ) are 0, 0.5, 1.6, 1.8 and 2.0 respectively</td>
</tr>
<tr>
<td>6.2</td>
<td>Wake firing rate ( (\lambda_1) ) steady states as a function of self-excitation parameter ( (\alpha) ) for ( \beta = 0.5 )</td>
</tr>
<tr>
<td>6.3</td>
<td>Sleep firing rate ( (\lambda_2) ) steady states as a function of self-excitation parameter ( (\alpha) ) for ( \beta = 0.5 )</td>
</tr>
<tr>
<td>6.4</td>
<td>Wake firing rate ( (\lambda_1) ) steady states as a function of mutual inhibition parameter ( (\beta) ) for ( \alpha = 0, 0.5, 1.6 )</td>
</tr>
<tr>
<td>6.5</td>
<td>Sleep firing rate ( (\lambda_2) ) steady states as a function of mutual inhibition parameter ( (\beta) ) for ( \alpha = 0, 0.5, 1.6 )</td>
</tr>
<tr>
<td>6.6</td>
<td>Wake firing rate ( (\lambda_1) ) steady states as a function of wake self-excitation ( (\alpha_{11}) ) for ( \beta = 0.5 ) and ( \alpha_{22} = 0.5, 1.6 )</td>
</tr>
<tr>
<td>6.7</td>
<td>Sleep firing rate ( (\lambda_2) ) steady states as a function of wake self-excitation ( (\alpha_{11}) ) for ( \beta = 0.5 ) and ( \alpha_{22} = 0.5, 1.6 )</td>
</tr>
<tr>
<td>7.1</td>
<td>Time course of ( \lambda_1 ) (in black) and ( \lambda_2 ) (in red); ( x )-axis: time in seconds, ( y )-axis: firing rate per second</td>
</tr>
</tbody>
</table>
7.2 Histograms for firing rates of the wake-active population($\lambda_1$) and the sleep active population($\lambda_2$) for $\alpha = 0$ ........................................... 68
7.3 Phase plane ($\lambda_1$, $\lambda_2$) for $\alpha = 0$ ........................................... 70
7.4 Time course of $\lambda_1$ (in black) and $\lambda_2$ (in red); $x$-axis: time in seconds, $y$-axis: firing rate per second ................................. 71
7.5 Histograms for firing rates of the wake-active population and the sleep active population for $\alpha = 0$ ........................................... 73
7.6 Phase plane ($\lambda_1$, $\lambda_2$) for $\alpha = 0$ ........................................... 73

8.1 Wake firing rate ($\lambda_1$) at steady state as a function of sleep firing rate ($\lambda_2$) for $\alpha_{13} = \alpha_{31} = 0.8$ and $\alpha_{11} = \alpha_{22} = 0, \beta_{12} = \beta_{32} = 0.5$ ........................................... 76
8.2 LC firing rate ($\lambda_3$) at steady state as a function of sleep firing rate ($\lambda_2$) for $\alpha_{13} = \alpha_{31} = 0.8$ and $\alpha_{11} = \alpha_{22} = 0, \beta_{12} = \beta_{32} = 0.5$ ........................................... 77
8.3 Wake firing rate ($\lambda_1$) at the unique steady state as a function of wake-promoting firing rate ($\lambda_3$) for $\beta = 0.5$ and $\alpha_{11} = \alpha_{22} = 0, \alpha_{13} = 0.8$ ........................................... 79
8.4 Sleep firing rate ($\lambda_2$) at the unique steady state as a function of wake-promoting firing rate ($\lambda_3$) for $\beta = 0.5$ and $\alpha_{11} = \alpha_{22} = 0, \alpha_{13} = 0.8$ ........................................... 80
8.5 Wake firing rate ($\lambda_1$) at the unique steady state as a function of wake-promoting firing rate ($\lambda_3$) for $\beta = 0.5$ and $\alpha_{11} = \alpha_{22} = 0.5, \alpha_{13} = 0.8$ ........................................... 81
8.6 Sleep firing rate ($\lambda_2$) at the unique steady state as a function of wake-promoting firing rate ($\lambda_3$) for $\beta = 0.5$ and $\alpha_{11} = \alpha_{22} = 0.5, \alpha_{13} = 0.8$ ........................................... 82
8.7 Wake firing rate ($\lambda_1$) at steady state as a function of $\alpha_{13}$ for $\beta = 0.5$, $\alpha_{31} = 0.5, \alpha_{11} = \alpha_{22} = 0$ and $s_1 = 20, s_2 = 10, s_3 = 10$ ........................................... 85
8.8 Sleep firing rate ($\lambda_2$) at steady state as a function of $\alpha_{13}$ for $\beta = 0.5$, $\alpha_{31} = 0.5, \alpha_{11} = \alpha_{22} = 0$ and $s_1 = 20, s_2 = 10, s_3 = 10$ ........................................... 85
8.9 LC firing rate ($\lambda_3$) at steady state as a function of $\alpha_{13}$ for $\beta = 0.5$, $\alpha_{31} = 0.5, \alpha_{11} = \alpha_{22} = 0$ and $s_1 = 20, s_2 = 10, s_3 = 10$ ........................................... 86
8.10 Wake firing rate ($\lambda_1$) at steady state as a function of $\alpha := \alpha_{13} = \alpha_{31}$ for $\beta = 0.5, \alpha_{11} = \alpha_{22} = 0$ and $s_1 = 20, s_2 = 10, s_3 = 10$ ........................................... 86
8.11 Sleep firing rate ($\lambda_2$) at steady state as a function of $\alpha := \alpha_{13} = \alpha_{31}$ for $\beta = 0.5$, $\alpha_{11} = \alpha_{22} = 0$ and $s_1 = 20, s_2 = 10, s_3 = 10$ ............... 87

8.12 LC firing rate ($\lambda_3$) at steady state as a function of $\alpha := \alpha_{13} = \alpha_{31}$ for $\beta = 0.5$, $\alpha_{11} = \alpha_{22} = 0$ and $s_1 = 20, s_2 = 10, s_3 = 10$ ............... 87

9.1 Time course of $\lambda_1$ (in black), $\lambda_2$ (in red) and $\lambda_3$ (in orange); $x$-axis: time in seconds, $y$-axis: firing rate per second ............... 89

9.2 Histograms for firing rates of the wake-active population (top two figures), the sleep active population, and the wake-promoting LC for $\alpha_{13} = \alpha_{31} = 0.6$, $\alpha_{11} = \alpha_{22} = 0$, $\beta = 0.5$ ............... 90

9.3 Survivor plots for sleep (left) and wake (right) on a semi-log scale (top) and a log-log scale (bottom) for $\alpha_{13} = \alpha_{31} = 0.4$, $\alpha_{11} = \alpha_{22} = 0$, $\beta = 0.5$ 92

9.4 Survivor plots for sleep (left) and wake (right) on a semi-log scale (top) and a log-log scale (bottom) for $\alpha_{13} = \alpha_{31} = 0.6$, $\alpha_{11} = \alpha_{22} = 0$, $\beta = 0.5$ 93

9.5 Survivor plots for sleep (left) and wake (right) on a semi-log scale (top) and a log-log scale (bottom) for $\alpha_{13} = \alpha_{31} = 0.8$, $\alpha_{11} = \alpha_{22} = 0$, $\beta = 0.5$ 94

11.1 C.c.d.f. plots for sleep (left) and wake (right) on a semi-log scale (top) and a log-log scale (bottom) for different ages *Plots from data courtesy of Blumberg et al. [Blu1]* ............... 104
# LIST OF TABLES

<table>
<thead>
<tr>
<th>TABLE</th>
<th>Description</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>7.1</td>
<td>Average values of the firing rates in the simulation for $\beta = 0.5$ and $\alpha = 0$</td>
<td>69</td>
</tr>
<tr>
<td>7.2</td>
<td>Average firing rates in simulation for $\beta = 0.5$ and $\alpha = 0.5$</td>
<td>72</td>
</tr>
<tr>
<td>7.3</td>
<td>Location of fixed points in the deterministic model</td>
<td>72</td>
</tr>
</tbody>
</table>
CHAPTER 1
INTRODUCTION

Several studies show that sleep is a universal phenomenon in mammalian species, and several non-mammalian species as well. The sleep times, durations, whether it is in one or two consolidated intervals every day or several short sleep intervals depends on the species, gender, age, habitat, climate, season, and many other factors. In other words, even though sleep itself is universal, the definition of sleep and its structure varies enormously. Some authors have attempted to speculate on the question of the function or functions of sleep, and why it is so strongly evolutionarily preserved across species. See for instance [Si].

The general approach however is to catalogue the different features of sleep structure in various species, the neural mechanisms and possible function in each species. One consequence of such studies would be to reveal which features are common across species and which are not. The times when an animal falls asleep and when the animal wakes up are random variables and so it is natural to ask about the properties of the probability distribution of these random variables. A feature that is found to be universal in adult mammals across species is that sleep bouts at all time scales are exponentially distributed and wake bouts at all time scales are distributed according to a power law.
I construct and analyze a mathematical model based on the known neurobiology in which the variables representing sleep and wake bout lengths follow the appropriate probability distributions. The model involves a combination of several different time scales,

1. 1 ms : Time scale of firing of action potential or spike in a single neuron

2. 1 s : Behavioral time scale, time scale over which states switch between wake and sleep

3. 1 day : Circadian variation, day/night variation in sleep

4. 2-3 weeks : Developmental time scale, age related changes
CHAPTER 2

BIOLOGICAL BACKGROUND AND DESCRIPTION OF THE BASIC PROBLEM

Different mammalian species and even different individuals within a species have distinct sleep patterns. During a consolidated sleep period, brief transitions into wake commonly occur across different species. Biologists ([Lo1], [Lo2], [Blu1]) are interested in studying this phenomenon in an attempt to understand the physiological basis and the function of sleep and these brief sleep-wake transitions.

Other aspects of sleep behavior such as circadian rhythms (rhythms due to time of the day) and homeostatic drive (‘need’ for sleep) have been studied more extensively both experimentally and in terms of mathematical modeling ([BA], [Ps]). These phenomena hold across time scales of several hours or days. As opposed to this, the brief sleep-wake transitions which we model in this thesis hold across a time scale of minutes, seconds or sometimes even shorter time scales. Moreover, the epochs at which these ‘brief awakenings’ occur and the time duration for which they last are random. This means that the relevant quantity of interest is not the precise moments of the beginning and end of these wake bouts or sleep bouts but rather the probability distributions of the wake and sleep bout durations.

Despite the many other differences across species, one common feature is that sleep
bouts in all mammalian species follow an exponential distribution while wake bouts follow a heavy-tailed power law. To be more precise, let $T_s$ be the random variable denoting the sleep bout duration and $T_w$ be the random variable denoting wake bout duration. Then it is observed that $P(T_s > t) \sim e^{-\kappa t}$ and $P(T_w > t) \sim t^{-k}$. The parameters $\kappa$ and $k$ in these distributions depend on the species, age of the animal and generally there are significant variations among individuals. There may also be slow changes in the values of these parameters depending on the time of the day and other factors. Nevertheless, it is striking that the form of the distribution is universal across species.

It has been noted ([Lo1], [Blu1]) that each bout is independent from another. When correlations at different scales are examined, there were no correlations found in the data either within wake or sleep bouts or among different type of bouts. This justifies pooling the bout data together to study the distribution without worrying about when in time the bout occurred.

Blumberg et al. [Blu1] studied these bout distributions in developing rats, i.e. in rats with age range between 2 days after birth and 3 weeks after birth. They observed that for rats aged 2 days, the bouts are exponentially distributed both for wake and sleep. As the animal develops, the sleep bouts retain the exponential distribution but the mean bout time increases with age. The wake bout distribution, on the other hand, develops a heavy tail. The content of the previous statement will be made more precise when we introduce the Generalized Pareto distribution to describe the wake bouts. For now, we simply note that at 21 days of age the wake bout distribution
is a power law with exponent close to 1 which is similar to what is found for adult animals.

Blumberg et al. also identify the populations of neurons in the brainstem of a developing rat which give rise to the sleep-wake transition behavior. They have hypothesized that a mutual inhibition mechanism is responsible for the switching behavior. Another finding is that a population of neurons called Locus Coeruleus is responsible for promoting wake activity and thus it may be involved in changing the wake bout distribution with age.

The problems addressed by this thesis and their solutions may be summarized as follows

1. Define a mathematical model based on the known physiology.

2. Analyse the model both numerically and mathematically. This includes exploring the parameter space of the model and classifying behavior.

3. Identify the mathematical mechanism that gives rise to power law distribution of the wake bouts, exponential distribution of the sleep bouts and the changes in the probability distribution that occur on time scale of several days or weeks.

4. Find a framework to describe wake and sleep bout distributions. The change in the bout distributions that occurs with age should be described by a smooth parameter. In particular, we need to describe the wake bouts using a probability distribution which has exponential and power law as special cases.
5. Find the best distribution fit to the data that is consistent with the data and 
the general framework described above.

6. Related to the previous two points, estimate distribution parameters from the 
data.

7. Show that the model is able to describe the properties of bout distribution 
deduced from the data.
CHAPTER 3
MODEL

3.1 Introduction

To model the basic circuit for the sleep-wake switch, we will consider as variables the sleep-active and wake-active populations. Each population of neurons (sleep-active or wake-active) is comprised of several hundred or several thousand neurons. In some cases, where neurons within a nucleus have extensive electrotonic coupling via gap junctions, the firing activity of all the neurons in the population is synchronous. In these cases, it makes sense to treat the entire nucleus or a connected sub-population as one entity and therefore model it using one dynamic variable. The Locus Coeruleus of developing mammals falls under this category. On the other hand, the other populations DLPT and PnO considered in this model are not comprised of neurons that fire synchronously. We will argue that it is reasonable even in these cases to treat the population as one variable if we were interested in modeling the brain states only.

The activity of individual neurons in an active population may vary widely. Some neurons may be quiescent while other individual neurons may have a high firing rate. Experimentally it is difficult to measure the activity of an individual neuron while
simultaneously recording population activity. We may only be able to infer individual neuron behavior based on population recordings. When we consider the population behavior, we find that there is a strong correlation between the firing activity and the behavioral state. So the labels of sleep-active and wake-active are indeed appropriate. Even though each individual wake-active neuron is not active at any given moment during wake period, the population activity is clearly elevated. The modulation in this firing activity within a state is usually negligible thus adding credence to the concept of discrete states.

The dynamical variable that we model is the firing rate of the population of neurons. Several authors have considered firing rate models, both in the deterministic and the stochastic setting. Different firing rate models tend to have different underlying assumptions. The main feature that is common to all such models is that the activity of either a single neuron or a population of neurons can be summarized using one variable i.e. the firing rate. The timings of the individual spikes and other details about the distribution of spikes are often ignored. In turn this feature depends on the following two main assumptions that the network properties are sufficiently explained by:

(a) The average response or firing rate of a population of neurons.

(b) The average effect of this firing rate on the input to the other population of neurons.

These assumptions in turn may be justified only if the firing is close to uniform and the typical time scale between spikes is several orders of magnitudes lower than the
time scale of the relevant phenomenon. In the case of the wake-active neurons in the Dorsolateral Pontine Tegmentum and the sleep-active neurons in the nucleus Pontis Oralis, the firing rate is usually between 1 to 10 Hz. This is comparable to the length of observable sleep and wake bouts which may be as low as 1 second. This means that the individual action potentials and their timing is an important consideration for the purposes of modeling sleep-wake transitions. Having said this, the firing rate essentially captures the behavioral state of the wake-active and the sleep-active nuclei, thus making the firing rate a very useful output variable.

In order to construct a mathematically tractable model of the phenomena while still making only assumptions that are biologically reasonable, we forego the second assumption mentioned above. We no longer assume that only the average activity of a population of neurons is ‘observed’ by another that receives inputs from the first. Each action potential has an inhibiting or exciting (or neutral) influence on the recipient nucleus. The effect of individual action potentials (as opposed to the long-time average) is particularly significant when the average firing rate is low i.e. particularly when the quiescent nucleus fires a rare action potential. In the extreme case, in a deterministic model, a quiescent nucleus will remain quiescent forever unless it receives external excitation. But in a stochastic model, random fluctuations can account for self-excitation of the quiescent nucleus till it reaches an active state.

The model thus has two components, the dynamical variable is the firing rate of the populations of neurons. This firing rate is affected by the activity of the other populations of neurons that are firing. We may visualize this as a graph where the vertices are the populations and the directed edges are the activity-based connections
The connections may be excitatory or inhibitory. In general, a connection doesn’t have to be uniformly excitatory or inhibitory nor does the coupling have to be constant but may be activity-dependent. However, for the purposes of our model, we will assume that the coupling is constant, the only (slow) modulation comes from the developmental time-scale where all couplings are being scaled upwards.

3.2 The main idea

We use a doubly stochastic Poisson process as a model for each population’s activity. These Poisson processes are coupled to the activity of the other populations of neurons. Specifically each population’s activity is modeled as a Poisson process with a firing rate $\lambda(t)$. The firing rate in turn is also a stochastic process which depends on the firing activity of other populations of neurons since these provide the input inhibitory or excitatory spikes.

For completeness and the convenience of the reader, we recall some basic definitions of a homogenous Poisson process, a non-homogenous Poisson process and a doubly stochastic Poisson process.

3.3 Definitions and Preliminaries

Definition 3.3.1. Given a set $T$ define a stochastic process $\{X(t)|t \in T\}$ to be a collection of random variables.
The set $T$ is called the index set of the process and in applications it is often interpreted as time. If $T$ is countable then the stochastic process is said to be a discrete-time process and if $T$ is an interval on $\mathbb{R}$, then the stochastic process is said to be a continuous-time process.

**Definition 3.3.2.** For a totally ordered set $T$, a non-decreasing process $\{N(t)|t \geq 0\}$ over the set of non-negative integers is known as a counting process.

A counting process counts the number of ‘events’ that have occurred in time $t$. For instance, a neuron firing action potentials can be modeled as a counting process where $N(t)$ would denote the number of action potentials fired by the neuron in time $t$.

**Definition 3.3.3.** A counting process is said to have independent increments if the number of events that occur in disjoint time intervals are independent.

**Definition 3.3.4.** A counting process has stationary increments if the distribution of the number of events that occur in any interval of time depends only on the length of the time interval i.e. if the distribution of the number of events on $(s, s + t)$ does not depend on $s$.

**Definition 3.3.5.** The counting process $\{N(t)|t \geq 0\}$ is called a Poisson process with rate $\lambda$, $\lambda > 0$ if

(i) $N(0) = 0$

(ii) $N(t)$ has independent increments.
(iii) The number of events in an interval of length $h$ has a Poisson distribution with mean $\lambda h$. i.e. for all $t, h \geq 0$

$$P(N(t + h) - N(t) = n) = e^{-\lambda h} \frac{(\lambda h)^n}{n!} \quad n = 0, 1, 2, ...$$

The following is an alternate definition of a Poisson process.

**Definition 3.3.6.** The counting process \{\(N(t)\mid t \geq 0\)\} is called a Poisson process with rate $\lambda$, $\lambda > 0$ if

(i) $N(0) = 0$

(ii) $N(t)$ has independent increments.

For all $t, h \geq 0$ the following two hold

(iii) $P(N(t + h) - N(t) = 1) = \lambda h + o(h)$

(iv) $P(N(t + h) - N(t) \geq 2) = o(h)$. 

**Theorem 3.3.7.** The above definitions are equivalent.

*Proof:* See, for instance, p.24-25 of [Kar].

An important generalization of the Poisson process is the nonhomogeneous Poisson process. The rate $\lambda$ is allowed to vary with time and the increments are not required to be stationary.
Definition 3.3.8. For a non-negative function $\lambda(t)$, the counting process $\{N(t)|t \geq 0\}$ is said to be a nonhomogeneous Poisson process with intensity function $\lambda(t), t \geq 0$ if the following are true

(i) $N(0) = 0$

(ii) $N(t)$ has independent increments.

(iii) $P(N(t + h) - N(t) = 1) = \lambda(t)h + o(h)$

(iv) $P(N(t + h) - N(t) \geq 2) = o(h)$.

A further generalization is obtained by making the intensity function $\lambda(t)$ into a random variable for each time $t$ or rather making $\lambda$ into a stochastic process over time $t$. To state it more appropriately, the Poisson process $N(t)$ has an intensity process $\lambda(t)$. Such a process has been studied starting with a seminal paper by Cox [Cox1]. A process where $N(t)$ is a nonhomogeneous Poisson process with an intensity process $\lambda(t)$ is also known as a doubly stochastic Poisson process or a Cox process.

We will use intensity function and intensity process interchangeably, although the latter is the more appropriate term.

Before giving the definition, we note that making the intensity $\lambda_i(t)$ of a Poisson process $N_i(t)$ depend on another Poisson process $N_j(t)$ where $j \neq i$ gives a doubly stochastic Poisson process.

Definition 3.3.9. Let $\{M(t)|t \geq t_0\}$ be a left-continuous, vector-valued stochastic
process. \( \{N(t)|t \geq t_0\} \) is a doubly stochastic Poisson process with intensity process \( \{\lambda(t,M(t))|t \geq t_0\} \) if for almost every given path \( \{\hat{M}(t)|t \geq t_0\} \) of the process \( \{M(t)|t \geq t_0\} \), \( N(t) \) is a Poisson process with intensity function \( \{\lambda(t,\hat{M}(t))|t \geq t_0\} \).

Sometimes, \( M(t) \) is referred to as the information process since it conveys information into \( N(t) \). For purposes of our model, for any given population \( i \), \( N_i(t) \) is the doubly stochastic Poisson process with intensity process \( \lambda_i(t,N(t)) \) where \( N(t) = (N_1(t), N_2(t), \ldots, N_n(t)) \).

With the preliminaries in place, we are ready to describe the model. The model consists of a system of ordinary stochastic first-order differential equations which give the dependence of the rate function \( \lambda_i \) on \( N(t) = (N_1(t), N_2(t), \ldots, N_n(t)) \).

### 3.4 SDE system

Suppose we consider \( n \) interacting populations. Let \( i \) be such that \( 1 \leq i \leq n \). The rate function \( \lambda_i(t) \) for the \( i \)th population is given by

\[
\lambda_i'(t) = f_i(\lambda_i) + \sum_{j=1}^n g_{ij}(\lambda_i) \sum_{k_j} \delta(T_{jk_j} - t) \tag{3.1}
\]

The model describes the time evolution of the intensity process \( \lambda_i \) for each population \( i \). Since the intensity is dependent on the random variables \( T_{jk_j} \), the intensity is not a function but a process (or a collection of random variables).

Note that \( \lambda_i \) is differentiable almost everywhere, i.e. except for points of jumps at
The interpretation of the notation $\lambda'_i$ at points of jump will be given in the next section. For now we note the following

1. For $1 \leq i, j \leq n$, $f_i$ and $g_{ij}$ are known functions of $\lambda_i$.

2. $\delta$ is the Dirac delta symbol.

3. $T_{jk}$ is a random variable denoting the times of occurrence of action potentials in the $j$th population.

4. For a fixed $i$ and $j$, $g_{ij}$ is either identically zero, or a strictly positive or a strictly negative function. The ‘signature’ of $g_{ij}$, that is whether it is zero, positive or negative, depends on the (directed) connection from $j$ to $i$.

   (a) If there is an excitatory connection from $j$ to $i$, then $g_{ij}$ is positive.

   (b) If there is an inhibitory connection from $j$ to $i$, then $g_{ij}$ is negative.

   (c) If there is no connection, then $g_{ij}$ is zero.

For the sake of compact notation, we will sometimes abbreviate the sum in the final term $\sum_{k_j} \delta(T_{jk} - t)$ and write it as $N'_j(t)$. This is merely for brevity, since we do not need to define the ‘derivative’ of $N_j$ for our purposes. As a heuristic explanation, we can think of $N_j$ as a sum of Heaviside functions with the location of ‘steps’ at random times $T_{jk}$. So $N'_j$ is just a collection of delta-functionals located at the random times $T_{jk}$. In the new notation the above equation becomes

$$\lambda'_i(t) = f_i(\lambda_i) + \sum_{j=1}^{n} g_{ij}(\lambda_i)N'_j(t) \quad (3.2)$$
3.5 Interpretation of the SDE system

1. Consider the simplest case of a single population which receives no inputs. In this case the second term in (3.1) is identically zero and so the equation for this population reduces to an ordinary uncoupled first order differential equation which can be solved by elementary means.

\[ \lambda'_i(t) = f_i(\lambda_i) \]  

(3.3)

Particularly important are the steady state solutions of the above differential equation. In absence of any external input, the firing rate of the neuron population will converge to a steady-state solution.

2. The second term in (3.1) is the interaction term. The summation is over all the neuron populations that are in the input. When the population \( j \) has a spike, the firing rate of \( i \) jumps by the amount \( g_{ij}(\lambda_i) \). If the connection is excitatory, the firing rate increases and if the connection is inhibitory the firing rate decreases.

In general, an action potential has the effect of increasing or decreasing the probability of firing in the post-synaptic neuron depending on whether the synaptic connection is excitatory or inhibitory respectively. The firing rate model captures the average effect of several such excitatory and inhibitory connections between two populations.
With this interpretation in mind, we explain the derivative term $\lambda'_i$. For simplicity, let’s assume $f_i(\lambda_i) = 0$ and also that there is only one $j$ in the input. The case where $f_i$ is not identically zero and there is more than one input makes the explanation more involved but is based on the same idea. The simplified expression is

$$\lambda'_i(t) = g_{ij}(\lambda_i)N'_j(t)$$

By the above equation, we simply mean the following

$$P(\lambda_i(t + h) - \lambda_i(t) = 0) = 1 - \lambda_j(t)h + o(h)$$
$$P(\lambda_i(t + h) - \lambda_i(t) = g_{ij}(\lambda_i(t))) = \lambda_j(t)h + o(h)$$
$$P(\lambda_i(t + h) - \lambda_i(t) \notin \{0, g_{ij}(\lambda_i(t))\}) = o(h) \quad (3.4)$$

The general case where $f_i$ is not identically zero and there is more than one input is a straightforward, even if slightly tedious generalization, which will be dealt with in section (4.1).

### 3.6 Graphical representation

The stochastic dynamical system described above can be captured in a graph where each vertex represents a single population of neurons. The directed arrows denote the direction of influence and each arrow has a positive or negative sign attached to it,
which indicates whether the connection is excitatory or inhibitory. To avoid clutter in the graphical representation, we use a regular arrow (→) to indicate an excitatory connection and a blunted arrow (⊣) to indicate an inhibitory connection.

In figure (3.1) we have a simple example of two vertices denoted $A$ and $B$ with an excitatory connection from $A$ to $B$ and an inhibitory connection from $B$ to $A$.

Other two commonly occurring motifs of mutually excitatory connection and mutually inhibitory connection are represented in figure (3.2).

### 3.7 Explicit forms of inhibition and excitation functions

For numerical investigation, we choose a particular form of the functions $f_i$ and $g_{ij}$.

1. $f_i(\lambda_i) = \frac{k_i - \lambda_i}{\tau_i}$, $k_i > 0$, $\tau_i > 0$

where $k_i$ is the ‘uncoupled’ firing rate of population $i$, i.e. $k_i$ is the unique steady-state firing rate of the population $i$ in absence of any external excitation or inhibition. $\tau_i$ is the ‘uncoupled’ relaxation time constant. Both these quantities
Figure 3.2: Commonly occurring motifs of mutual excitation and mutual inhibition can be measured in principle. This term $f_i(\lambda_i)$ can be thought of as the restoring term since it ensures that in absence of any other excitation or inhibition, $f_i$ ensures that the firing rate returns to its uncoupled steady state.

2. When the coupling from population $j$ to population $i$ is inhibitory, the coupling coefficient $g_{ij}$ is given by

$$g_{ij}(\lambda_i) = -\beta_{ij}\lambda_i, \quad 0 \leq \beta_{ij} \leq 1$$

where $\beta_{ij}$ is the inhibition constant for the coupling from population $j$ to population $i$. In this case, $g_{ij}$ will be referred to as the inhibition coefficient. So in particular, we note that the inhibition coefficient is linear in $\lambda_i$.

Note that $\beta_{ij}$ has the upper bound of 1. This is to avoid the possibility of
an inhibitory input resulting in a negative firing rate. The reason $\beta_{ij}$ is non-negative is so that the inhibition term does not act as excitation.

3. When the coupling from population $j$ to population $i$ is excitatory, the coupling coefficient is

$$g_{ij}(\lambda_i) = \alpha_{ij} \lambda_i \left( 1 - \frac{\lambda_i}{s_i} \right)$$

where $0 \leq \alpha_{ij} \leq 1$, $s_i > k_i > 0$, $0 \leq \lambda_i \leq s_i$

where $\alpha_{ij}$ is the excitation constant for the coupling from population $j$ to population $i$. For an excitatory connection, $g_{ij}$ will be referred to as the excitation coefficient. The excitation coefficient, unlike the inhibition coefficient is quadratic in $\lambda_i$. In fact, the excitation coefficient is modeled as a logistic growth term. For small values of $\lambda_i$, the excitation coefficient is proportional to $\lambda_i$. But as the firing rate $\lambda_i$ increases, many or most of the neurons in the population start firing. Each individual neuron has a refractory period after a spike occurs, during which another spike does not occur or occurs with very low probability. This in turn implies that the population firing rate can also not increase indefinitely. In fact when $\lambda_i$ is close to saturation or firing close to its maximum possible rate, any new excitatory input will have very little effect on increase in the firing rate. These properties of the biological system are modeled using the quadratic logistic term for the excitation coefficient. We note that we
have restricted the domain of $\lambda_i$, $s_i$ is the upper bound which ensures that $g_{ij}$ is a strictly non-negative function.

The fact that $\lambda_i$ is quadratic is also significant when we do bifurcation analysis. We will show that this leads to multiple stable steady-state solutions of the differential equations. This multi-stability has the consequence of changing the firing rate distribution of all cell populations which in turn leads to a change in the distribution of wake and sleep bouts.

The excitation constant $\alpha_{ij}$ is restricted to be a number between 0 and 1. Clearly we need $\alpha_{ij} > 0$, otherwise the excitation term will act as an inhibition. The reason why we need an upper bound on $\alpha_{ij}$ is provided by the following theorem.

**Theorem 3.7.1.** For all $t \geq 0$, $\lambda_i(t) \leq s_i$ if and only if $\alpha_{ij} \leq 1$.

**Proof:** By hypothesis, for a Poisson process two jumps cannot occur simultaneously. So we require that the magnitude of a single jump should not be so large as to make the firing rate $\lambda_i$ exceed the supremum of $s_i$. In other words, we require that

$$\lambda_i + \alpha_{ij} \lambda_i \left(1 - \frac{\lambda_i}{s_i}\right) \leq s_i$$

for $0 \leq \lambda_i \leq s_i$. If $\alpha_{ij} = 0$, then the magnitude of the jump is 0. So we will assume without loss of generality that $\alpha_{ij} > 0$.

If we let $x := \frac{\lambda_i}{s_i}$ then $x \in [0, 1]$. Let $\alpha := \alpha_{ij}$. So we require that

$$x(1 + \alpha - \alpha x) \leq 1 \quad \forall x \in [0, 1]$$

21
or in other words the requirement that $\forall t > 0$, $\lambda_i(t) \leq s_i$ is equivalent to the following

$$\max_{x \in [0,1]} f(x) \leq 1$$

(3.5)

where $f(x) := x(1 + \alpha - \alpha x)$.

Note that $f'(x) = 1 + \alpha - 2\alpha x$ and $f''(x) = -2\alpha \leq 0$. The maximum of $f(x)$ is reached where $f'(x) = 0$ or at the end-points 0 and 1. We note that $f'(x^*) = 0 \iff x^* = \frac{1 + \alpha}{2\alpha} = \frac{1}{2} \left(1 + \frac{1}{\alpha}\right)$.

Suppose that $\alpha > 1$. Then $\frac{1}{2} \left(1 + \frac{1}{\alpha}\right) \in (0, 1)$. This means that

$$\max_{x \in [0,1]} f(x) = f\left(\frac{1}{2} \left(1 + \frac{1}{\alpha}\right)\right)$$

$$= \frac{1 + \alpha}{2\alpha} \left(1 + \alpha \left(1 - \frac{1 + \alpha}{2\alpha}\right)\right)$$

$$= \frac{(1 + \alpha)^2}{4\alpha} > 1$$

(3.6)

where we used $\alpha > 1 \Rightarrow (1 - \alpha)^2 > 0 \Rightarrow (1 + \alpha)^2 > 4\alpha \Rightarrow \frac{(1 + \alpha)^2}{4\alpha} > 1$ in the final inequality.

But the inequality (3.6) contradicts the requirement (3.5) and so we conclude that $\alpha$ cannot be greater than 1.

Now assume that $0 < \alpha \leq 1$. We note that $\frac{1}{2} \left(1 + \frac{1}{\alpha}\right)$ is a decreasing function of $\alpha$ and $\frac{1}{2} \left(1 + \frac{1}{\alpha}\right) \to \infty$ as $\alpha \to 0$ and $\frac{1}{2} \left(1 + \frac{1}{\alpha}\right) = 1$ when $\alpha = 1$. So
for $x \in [0, 1]$, $x \leq \frac{1}{2} (1 + \frac{1}{\alpha})$ for all $0 < \alpha \leq 1$. This means that $f'(x) = 1 + \alpha - 2\alpha x \geq 1 + \alpha - 2\alpha \left(\frac{1 + \alpha}{2\alpha}\right) = 0$ and so

$$\max_{x \in [0,1]} f(x) = f(1) = 1$$

which satisfies condition (3.5). This proves the theorem. □

Note: We note that the parameters $k_i$ and $s_i$ have the dimensions of rate or time$^{-1}$, $\tau$ has the dimensions of time and $\alpha_{ij}$ and $\beta_{ij}$ are dimensionless.

### 3.8 Circuit for sleep-wake switch

We recapitulate some important biological details.

The main component of the sleep-wake switch is comprised of a wake-active population (Dorsolateral Pontine Tegmentum or DLPT for short) and a sleep-active population (nucleus Pontis Oralis or PnO for short) with a mutually inhibitory connection between the two. Mutual inhibition has been proposed for a switch-like mechanism in other parts of the nervous system. A deterministic model in absence of external inputs with merely mutual inhibition is insufficient to make the switching mechanism work. However a stochastic switch that uses mutual inhibition as its main component is indeed possible as shown by our modeling study and supported by experimental data on DLPT and PnO and the interaction between the two populations. Even in the stochastic model, external inputs have an auxiliary role such as stabilizing the switch or providing circadian variation.
Experiments in adults have shown that Locus Coeruleus (LC) has an important role to play in wake-promoting behavior. In the case of infants too, LC has been shown to modulate wake behavior by Blumberg et al [Blu1]. The modulation however is complex and seems to have a dual effect. On one end the presence of LC and a strong connection with the wake-active population increases the proportion of longer wake bouts and on the other end, the same presence of LC seems to be to make the shorter bouts terminate even earlier. A hypothesis that has been proposed for this dual effect is that natural selection acts to conserve the total amount of sleep time and therefore
total amount of wake time. If changes in the developing circuit lead naturally to longer wake bouts, the extra time might come at the expense of the shorter wake bouts in order to maintain the same sleep time.

Even though such a hypothesis may well be true, it is not clear that what the mechanism might be. It may well be that a single mechanistic explanation is able to account for the changes in wake bouts at both ends of the spectrum. Alternately, it may be that naturally occurring changes at one end of the spectrum set up a strong selection pressure over an evolutionary time scale for changes to occur at the other end of the spectrum, eventually leading the two effect to be synchronized. Our model does not have the complexity needed to tell apart these two alternative scenarios. But we hope to return to this question in future, more elaborate modeling studies.

The precise connection of LC to the basic sleep-wake switch is not clear yet. However, we postulate a specific mutual excitatory connection between LC and DLPT based on our modeling study. We also postulate an inhibitory connection from PnO to LC. It is probably the case that there are many other direct and indirect connections between these nuclei and also there are modulating effects of other nuclei not accounted for. But as we will see, the basic model of mutual inhibition between DLPT and PnO and mutual excitation between LC and DLPT is able to account for most of the important features in the data.

We summarize the main qualitative ingredients of our model.

1. We focus on three neuronal populations which are labelled wake-active, sleep-active and wake-promoting populations. In equations, we often use subscripts
to refer to these populations and it is convenient to use subscripts 1, 2 and 3 respectively for the three populations.

2. Both the wake-active and sleep-active populations have self-excitation which is transmitted through excitatory action potentials.

3. The wake-active and sleep-active populations are connected to each other via a mutual inhibition.

4. The wake-active and wake-promoting populations are mutually excitatory.

All these features are neatly summarized in the graphical representation in figure (3.3). We can directly write down the system of stochastic differential equations from this graphical representation.

The dynamical system corresponding to the above diagram is

\[
\begin{align*}
\lambda_1' &= \frac{k_1 - \lambda_1}{\tau_1} - \beta_{12} \lambda_1 N_2' + \alpha_{11} \lambda_1 \left( 1 - \frac{\lambda_1}{s_1} \right) N_1' + \alpha_{13} \lambda_1 \left( 1 - \frac{\lambda_1}{s_1} \right) N_3' \\
\lambda_2' &= \frac{k_2 - \lambda_2}{\tau_2} - \beta_{21} \lambda_2 N_1' + \alpha_{22} \lambda_2 \left( 1 - \frac{\lambda_2}{s_2} \right) N_2' \\
\lambda_3' &= \frac{k_3 - \lambda_3}{\tau_3} - \beta_{32} \lambda_3 N_2' + \alpha_{31} \lambda_3 \left( 1 - \frac{\lambda_3}{s_3} \right) N_1'
\end{align*}
\]  

(3.7)

with the initial conditions \(0 < \lambda_i(0) < s_i\) and the parameter values \(\tau_i > 0, s_i > k_i > 0\) and \(0 \leq \alpha_{ij}, \beta_{ij} \leq 1\). Unless mentioned otherwise, we will always assume that these parameter values and initial conditions hold true.
Here the subscript 1 is for DLPT or wake-active, subscript 2 for PnO or sleep-active and subscript 3 for LC or wake-promoting populations. For young infants age postnatal day 2 for instance, LC does not play a significant role and so the above system reduces to

\[
\lambda_1' = \frac{k_1 - \lambda_1}{\tau_1} - \beta_{12}\lambda_1 N_2' + \alpha_{11}\lambda_1 \left(1 - \frac{\lambda_1}{s_1}\right) N_1'
\]

\[
\lambda_2' = \frac{k_2 - \lambda_2}{\tau_2} - \beta_{21}\lambda_2 N_1' + \alpha_{22}\lambda_2 \left(1 - \frac{\lambda_2}{s_2}\right) N_2'
\]

The above systems (3.7) and (3.8) are suitable for numerical analysis. Before we describe the results of numerical simulation and comparison with data, we can get some further insight by suitable mathematical analysis. We will show that the stochastic system presented above may be approximated by a deterministic system. The dynamical system thus obtained can then be studied with the help of results from bifurcation analysis. The steady-state solutions can be easily obtained numerically. We will show that the steady-state solutions of this deterministic system can tell us about the behavior of the stochastic system.
CHAPTER 4
DETERMINISTIC SYSTEM

4.1 Derivation of the deterministic system

In this section, we derive a system of deterministic first-order ordinary differential equations which may be thought of as a ‘first-order’ approximation to the stochastic system that we are studying. This is because the expected value of the collection of intensity processes \( \{\lambda_i, 1 \leq i \leq n\} \) obeys this deterministic dynamical system for a sufficiently small time. We make this statement more precise in this section.

The definition (3.3.9) of a doubly stochastic Poisson process given earlier is not suitable for purposes of calculations. We first make some observations that will help us apply this definition.

**Definition 4.1.1.** Let \( \Lambda \) denote the collection of the intensity processes \( \{\lambda_i, 1 \leq i \leq n\} \). We will refer to \( \Lambda \) as the complete intensity process.

**Definition 4.1.2.** For a given subset \( S \) of the index set \( T \), we refer to the realization of the vector-valued random variables \( \{\Lambda(t)|t \in S\} \) as a realization or a sample path of the complete intensity process \( \Lambda \). This sample path will be denoted using a subscript i.e. \( \Lambda_S = \{\Lambda(t)|t \in S\} \)
Note: In particular, the value of the complete intensity process at time \( t_0 \) is denoted by \( \Lambda_{t_0} \).

If \( t_0 \) is the current time, then presumably the process upto the current time \( t_0 \) has been either recorded experimentally or is presumed to be known otherwise in the sense that a particular sample path has been fixed upto the present time.

**Definition 4.1.3.** The observed sample path up to the present time \( t_0 \) will be referred to as the history of the process and denoted by \( \mathcal{H}_{t_0} := \{ \Lambda_t \mid -\infty \leq t \leq t_0 \} = \Lambda_{(-\infty \leq t \leq t_0)} \).

so that \( \Lambda_t \) is a particular point on the sample path and the future evolution of \( \Lambda \) for \( t \geq t_0 \) is conditional on \( \mathcal{H}_{t_0} \). In fact in our model, the future evolution of \( \{ \Lambda(t) \mid t \geq t_0 \} \) only depends on \( \Lambda_{t_0} \).

**Theorem 4.1.4.** The complete intensity process \( \Lambda \) is a Markov process.

**Proof:** All future probabilities are conditioned on the observed history of the process. For each \( i \), \( N_i \) is a Poisson process with the intensity process \( \lambda_i \). If \( t \) is the present time then conditioned on \( \mathcal{H}_t \), \( N_i \) is a Poisson process with the intensity function \( \lambda_i \) and so the following must hold

\[
P(N_i(t, t+h) = 0 | \mathcal{H}_t) = 1 - \lambda_i(t)h + o(h) \quad (4.1)
\]
\[
P(N_i(t, t+h) = 1 | \mathcal{H}_t) = \lambda_i(t)h + o(h) \quad (4.2)
\]
\[
P(N_i(t, t+h) > 1 | \mathcal{H}_t) = o(h) \quad (4.3)
\]

29
However, the right hand side of the above equations does not depend on firing rates in the past that is, firing rates before time $t$ and so we may replace $\mathcal{H}_t$ above by $\Lambda_t$ and get

\begin{align*}
P(N_i(t, t + h) = 0|\Lambda_t) &= 1 - \lambda_i(t)h + o(h) \quad (4.4) \\
P(N_i(t, t + h) = 1|\Lambda_t) &= \lambda_i(t)h + o(h) \quad (4.5) \\
P(N_i(t, t + h) > 1|\Lambda_t) &= o(h) \quad (4.6)
\end{align*}

In turn, $\Lambda$ depends on $\{N_i|1 \leq i \leq n\}$ only through the system (3.1) or through (3.2). This shows that $\Lambda$ is a Markov process. □

Since $N_i$ is a Poisson process with intensity $\lambda_i(t, M_i(t))$ or simply $\lambda_i$, it must be true for all $i$ that

$$\lambda_i(t) = \lim_{h \to 0^+} \frac{P(N_i(t, t + h) > 0)}{h}$$

and also that

$$P(N_i(t, t + h) > 1) = o(h)$$

In fact we require a slightly stronger condition.

$$P \left( \sum_{i=1}^{n} N_i(t, t + h) > 1 \right) = o(h) \quad (4.7)$$

that is, we require that two spikes do not occur simultaneously not just within a population but also among all populations. This condition will ensure that the input
to a given population of neurons does not change suddenly in a short time interval and also that the recipient population does not receive two simultaneous inputs from two different inputs.

**Lemma 4.1.5.** Suppose $\lambda$ is the firing rate for a population that receives no inputs. For a sufficiently smooth function $f$, for instance for $f \in C_1$, 

$$\lambda(t + h) = \lambda(t) + hf(\lambda) + o(h)$$

**Proof:** Since $f$ satisfies the differential equation $\lambda' = f(\lambda)$, we have 

$$\lambda(t + h) = \lambda(t) + h\lambda'(t) + o(h)$$

from which the result follows. □

We want to know what is the expected change in the firing rate of $\lambda_i$ over a sufficiently small time $h$. When there are no inputs, it is given by the above lemma. When there are inputs $\lambda_i$ follows the differential equation 

$$\lambda'_i = f_i(\lambda_i) + \sum_{j=1}^{n} g_{ij}(\lambda_i)N_j'$$

For notational ease we drop the subscript $i$, and so

$$\lambda' = f(\lambda) + \sum_{j} g_j(\lambda)N_j'$$

We now state the main theorem of this section

**Theorem 4.1.6.** The expected change in firing rate in time $h$ is given by

$$E[\lambda(t + h) - \lambda(t) | \Lambda_t] = hf(\lambda) + \sum_{j} hg_j(\lambda(t))\lambda_j(t) + o(h) \quad (4.8)$$
Proof: In the following calculations, all random variables and random processes are conditioned on \( \Lambda_t \).

The condition (4.7) ensures that more than one spike does not occur in a small time \( h \) and also that the input \( \lambda_j \) is not rapidly changing in this time. So for any \( j \) in the input we may write using Lemma (4.1.5)

\[
\lambda(t + h) = \lambda(t) + hf(\lambda) + g_j(\lambda(T_j)) + o(h) \quad \text{with probability} \quad \lambda_j(t)h + o(h)
\]

The above is based on the probability of one spike occurring in the time interval \([t, t + h]\). \( T_j \) is the random variable denoting the time of occurrence of the spike and so \( T_j \in [t, t + h] \). Let \( T_j = t + \epsilon h \) for some \( \epsilon \in [0, 1] \) and so

\[
\lambda(t + h) = \lambda(t) + hf(\lambda) + g_j(\lambda(t + \epsilon h)) + o(h) \quad \text{with probability} \quad \lambda_j(t)h + o(h) \quad (4.9)
\]

Since in the absence of spikes \( \lambda \) is a smooth function of time,

\[
\lambda(t + \epsilon h) = \lambda(t) + \epsilon h \lambda'(t) + o(h)
\]

and so since \( g_j \) is a smooth function of \( \lambda \),

\[
g_j(\lambda(t + \epsilon h)) = g_j(\lambda(t) + \epsilon h \lambda'(t) + o(h)) = g_j(\lambda(t)) + (\epsilon h \lambda'(t) + o(h))g_j'(\lambda(t))
\]

and so

\[
g_j(\lambda(t + \epsilon h))h = hg_j(\lambda(t)) + o(h) \quad (4.10)
\]

Using the probability that no spike occurs in the input in the time \( h \), we may write

\[
\lambda(t + h) = \lambda(t) + hf(\lambda) + o(h) \quad \text{with probability} \quad 1 - \sum_j \lambda_j(t)h + o(h) \quad (4.11)
\]
and so the expected value of $\lambda(t + h)$ is given by

$$
E[\lambda(t + h) | \Lambda_t] - \lambda(t) = \left( hf(\lambda) + o(h) \right) \left( 1 - \sum_j \lambda_j(t) h + o(h) \right) 
+ \sum_j \left( hf(\lambda) + g_j(\lambda(t + \epsilon h)) + o(h) \right) \left( \lambda_j(t) h + o(h) \right) + o(h)
$$

(4.12)

Using (4.10), we simplify the above

$$
E[\lambda(t + h) | \Lambda_t] - \lambda(t) = hf(\lambda) + \sum_j h g_j(\lambda(t)) \lambda_j(t) + o(h)
$$

(4.13)

which proves the result. □

**Theorem 4.1.7.**

$E[\lambda'(t) | \Lambda_t] = f(\lambda(t)) + \sum_{j \in \text{input}} g_j(\lambda(t)) \lambda_j(t)$

(4.14)

*Proof:* From the previous theorem

$$
E \left[ \frac{\lambda(t + h) - \lambda(t)}{h} \right] | \Lambda_t = f(\lambda) + \sum_j h g_j(\lambda(t)) \lambda_j(t) + \frac{o(h)}{h}
$$

(4.15)

and taking the limit $h \to 0$, we get that

$$
E[\lambda'(t) | \Lambda_t] = f(\lambda(t)) + \sum_{j \in \text{input}} g_j(\lambda(t)) \lambda_j(t)
$$

(4.16)

□

**Theorem 4.1.8.**

$$
\lim_{h \to 0} \left[ \frac{Var[\lambda(t + h) - \lambda(t)]}{h} \right] | \Lambda_t = \sum_j (g_j(\lambda(t)))^2 \lambda_j(t)
$$

(4.17)
Proof: First from Theorem (4.1.6), it is easy to see that

\[ E^2[\lambda(t + h) - \lambda(t)]|\Lambda_t] = o(h) \]  

(4.18)

From equations (4.9) and (4.11),

\[
E[(\lambda(t + h) - \lambda(t))^2|\Lambda_t] = \left( hf(\lambda) + o(h) \right)^2 \left( 1 - \sum_j \lambda_j(t)h + o(h) \right) 
\]

\[
+ \sum_j \left( hf(\lambda) + g_j(\lambda(t + \epsilon h)) + o(h) \right)^2 \left( \lambda_j(t)h + o(h) \right) + o(h) \]  

(4.19)

and so

\[
E[(\lambda(t + h) - \lambda(t))^2|\Lambda_t] = \sum_j g_j^2(\lambda(t + \epsilon h))\lambda_j(t)h + o(h) 
\]

\[
= \sum_j g_j(\lambda(t + \epsilon h)) \left( hg_j(\lambda(t)) + o(h) \right)\lambda_j(t) + o(h) 
\]

\[
= \sum_j hg_j^2(\lambda(t))\lambda_j(t) + o(h) 
\]  

(4.20)

where we used (4.10) in the last two steps.

Finally, using (4.18) and (4.20), we get that

\[
Var[\lambda(t + h) - \lambda(t)|\Lambda_t] = \sum_j h(g_j(\lambda(t)))^2\lambda_j(t) + o(h) 
\]  

(4.21)

from which the result follows. □

Before applying the theorem to the sleep-wake system, we should mention recent work by Rodriguez et al. [Ro1] and [Ro2] where the authors consider a different kind of a stochastic system, specifically a system of diffusion equations, and then consideration of various moments leads to deterministic differential equations. The theory of stochastic bifurcation is fairly new and an application can be found in [Xu] where stochastic bifurcation in Duffing system is studied.
4.2 Application of Results to the Sleep-Wake Model

Now we apply Theorem (4.14) to the model of sleep-wake cycling to get the deterministic dynamical system for the firing rate functions.

Using the compact notation \( \tilde{\lambda} = E[\lambda(t)|\Lambda_t] \) and \( \tilde{\lambda}' = E[\lambda'(t)|\Lambda_t] \), we may write the above equation as

\[
\tilde{\lambda}'(t) = f(\tilde{\lambda}(t)) + \sum_{j \in \text{input}} g_j(\tilde{\lambda}(t))\tilde{\lambda}_j(t)
\]

Finally, dropping the tildes and including subscripts for the population, we get the following dynamical system

\[
\lambda'(t) = f(\lambda(t)) + \sum_{j=1}^{n} g_j(\lambda(t))\lambda_j(t)
\]

Explicitly the three population model for wake-sleep switch becomes

\[
\begin{align*}
\lambda_1' &= \frac{k_1 - \lambda_1}{\tau_1} - \beta_{12}\lambda_1\lambda_2 + \alpha_{11}\lambda_1^2 \left(1 - \frac{\lambda_1}{s_1}\right) + \alpha_{13}\lambda_1 \left(1 - \frac{\lambda_1}{s_1}\right) \lambda_3 \\
\lambda_2' &= \frac{k_2 - \lambda_2}{\tau_2} - \beta_{21}\lambda_2\lambda_1 + \alpha_{22}\lambda_2^2 \left(1 - \frac{\lambda_2}{s_2}\right) \\
\lambda_3' &= \frac{k_3 - \lambda_3}{\tau_3} - \beta_{32}\lambda_3\lambda_2 + \alpha_{31}\lambda_3 \left(1 - \frac{\lambda_3}{s_3}\right) \lambda_1
\end{align*}
\]

with the initial conditions \( 0 < \lambda_i(0) < s_i \) and the parameter values \( \tau_i > 0, s_i > k_i > 0 \) and \( 0 \leq \alpha_{ij}, \beta_{ij} \leq 1 \). In fact, if we were only interested in the deterministic system without reference to the original stochastic differential equations, we could relax the restrictions on the values of the excitation and inhibition parameters and let \( \alpha_{ij}, \beta_{ij} \geq 0 \). We will often do so for ease of analysis. For instance when following
a bifurcation curve which has a bifurcation for a value of say $\alpha_{ij} > 1$, it is easier to understand the qualitative behavior if we first include all non-negative values of $\alpha_{ij}$ and then eventually suitably restrict these values.

To summarize, unless otherwise mentioned, we have the following parameter values for the deterministic model

$$
\tau_i > 0 \\
\delta_i > k_i > 0 \\
\alpha_{ij}, \beta_{ij} \geq 0 \quad (4.25)
$$

Let $\Lambda = (\lambda_1, \lambda_2, \ldots, \lambda_n)$. For $i \in \{1, 2, 3\}$ we write the above more compactly as

$$
\lambda'_i = h_i(\Lambda) \quad (4.26)
$$
or in vector notation as

$$
\Lambda' = h(\Lambda) \quad (4.27)
$$
CHAPTER 5
RULING OUT CLOSED ORBITS AND CHAOTIC BEHAVIOR

Before a detailed analysis of the deterministic system, we can rule out a huge class of solutions at one stroke. The system (4.24) that we have constructed is a strongly monotone system. A strongly monotone system has the following two properties

1. Any subsystem is also strongly monotone.

2. Except for a measure-zero set, trajectories with arbitrary initial conditions converge to fixed point solutions.

We can thus rule out periodic orbits or chaotic behavior right away. It has been hypothesized that many biological systems or subsystems may evolve to be monotone systems because of the desirable properties of monotone systems such robustness and behavior that is almost like a one-dimensional system. A suitable characterization of a strongly monotone system and a more detailed explanation of what makes the sleep-wake system monotone is given in Appendix B.

Here we give a simple proof of the fact that a two-dimensional system with mutual inhibition does not permit closed orbits. Of course the full theorem (B.0.4) says much more and the result is only non-trivial in higher dimensions. However, the two
dimensional case provides some insight into the result. By $x$ inhibiting $y$, we mean that an instantaneous increase in the value of $x$ causes $y'$ to become negative or causes $y$ to decrease as a function of time. When we say that such an inhibition is monotone, we mean that the inhibitory effect does not depend on the particular parameter or variable values. We state this more precisely in the following definition.

**Definition 5.0.1.** More precisely, $x$ inhibits $y$ monotonously if and only if for all $x_1 > x_2$ we have that $\frac{dy}{dt}(x_1, y) < \frac{dy}{dt}(x_2, y)$ for all $y \in \mathbb{R}$.

**Theorem 5.0.2.** For $(x, y) \in \mathbb{R}^2$ let $x' = f(x, y)$ and $y' = g(x, y)$ be a monotone mutually inhibitory system. Then the system does not have any closed orbit solutions.

**Proof:** Suppose there is a closed orbit solution. We can find two points on this orbit with the same $y$-coordinate, say $(x_1, y)$ and $(x_2, y)$ such that $x_2 > x_1$. By the hypothesis of monotone inhibition of $x$ on $y$, this implies that $g(x_1, y) > g(x_2, y)$. For a closed orbit $g(x_1, y)$ and $g(x_2, y)$ have opposite signs, and so the only consistent sign assignment is $g(x_1, y) > 0$ and $g(x_2, y) < 0$, which rules out a counter-clockwise closed orbit.

On a closed orbit, we can find two points with the same $x$-coordinate, say $(x, y_1)$ and $(x, y_2)$ with $y_2 > y_1$. Since $y$ inhibits $x$ monotonously, $f(x, y_1) > f(x, y_2)$. So the only consistent sign assignment is $f(x, y_1) > 0$ and $f(x, y_2) < 0$, which rules out clockwise closed orbits.

Thus there are no closed orbit solutions. □
CHAPTER 6
ANALYSIS OF THE TWO COMPONENT DETERMINISTIC SYSTEM

Before analyzing the full system, let’s consider the reduced system consisting of wake-active and sleep-active populations only. The reason for studying the reduced system is not merely for ease of analysis, this two component system is responsible for wake-sleep switching for early infants of age post-natal day 2, for example. The wake-promoting population Locus Coeruleus becomes active only towards the end of the first post-natal week. Thus when the excitation parameters between the wake-promoting and wake-active populations are tuned down, the three component system should reduce to the two component system and we get

\[ \begin{align*}
\lambda_1' &= \frac{k_1 - \lambda_1}{\tau_1} - \beta_{12} \lambda_1 \lambda_2 + \alpha_{11} \lambda_1^2 \left( 1 - \frac{\lambda_1}{s_1} \right) \\
\lambda_2' &= \frac{k_2 - \lambda_2}{\tau_2} - \beta_{21} \lambda_2 \lambda_1 + \alpha_{22} \lambda_2^2 \left( 1 - \frac{\lambda_2}{s_2} \right)
\end{align*} \]  

(6.1)

where \( \alpha_{11}, \alpha_{22} \geq 0, \beta_{12}, \beta_{21} \geq 0, \tau_1, \tau_2 > 0, s_1 > k_1 > 0 \) and \( s_2 > k_2 > 0 \).

The first term on the right hand side is the autonomous restoring term, the second term is the inhibition term and the third term is the self-excitation term. Except
for the parameters of autonomous firing rate, the inhibition parameters and the self-
excitation parameters, there is an obvious symmetry between the wake-active pop-
ulation 1 and the sleep-active population 2. This is more apparent in the graphical
representation. The symmetry between the wake-active and sleep-active populations
is broken when the wake-promoting population Locus Coeruleus becomes active. This
is the reason for the difference in bout distributions of wake and sleep.

6.1 Reduced system with zero self-excitation

First we consider the case when the self-excitation parameters $\alpha_{11}$ and $\alpha_{22}$ for the
two populations are negligible. For $\alpha_{11} = \alpha_{22} = 0$, the above system reduces to

$$
\lambda'_1 = \frac{k_1 - \lambda_1}{\tau_1} - \beta_{12}\lambda_1\lambda_2
$$

$$
\lambda'_2 = \frac{k_2 - \lambda_2}{\tau_2} - \beta_{21}\lambda_2\lambda_1
$$

(6.2)

**Definition 6.1.1.** For a dynamical system with a dynamical variable $\lambda$, define the $\lambda$
-nullcline to be the curve defined by $\lambda' = 0$.

**Definition 6.1.2.** The Jacobian of the system $\Lambda' = h(\Lambda)$ is the matrix

$$
J(h(\Lambda)) = \begin{bmatrix}
\frac{\partial h_i}{\partial \lambda_j}
\end{bmatrix}_{1 \leq i,j \leq n}
$$

**Theorem 6.1.3.** (Condition for stability) A steady state solution of $\Lambda' = h(\Lambda)$ i.e.
a solution of $h(\Lambda) = 0$ is asymptotically stable if all the eigenvalues of the Jacobian
$J(h(\Lambda))$ are strictly negative.
Proof: See for instance [Guc] or [Wig]. □

The nullclines for the above system (6.2) are given by

\[
\begin{align*}
\frac{k_1 - \lambda_1}{\tau_1} - \beta_{12}\lambda_1\lambda_2 &= 0 \\
\frac{k_2 - \lambda_2}{\tau_2} - \beta_{21}\lambda_2\lambda_1 &= 0
\end{align*}
\]  (6.3)

from which it is easy to see that for

\[
\lambda_1 = \frac{k_1}{1 + \beta_{12}\tau_1\lambda_2} \quad (6.4)
\]
\[
\lambda_2 = \frac{k_2}{1 + \beta_{21}\tau_2\lambda_1} \quad (6.5)
\]

Note that all the parameters \( \beta_{ij}, \tau_i \) and \( k_i \) are non-negative and the rates \( \lambda_i \) are also non-negative variables.

**Definition 6.1.4.** By non-negative (positive) steady state we mean a steady state for which all its coordinates are non-negative (positive).

**Theorem 6.1.5.** For all \( \beta_{ij} \geq 0, \tau_i > 0 \) and \( k_i > 0 \) where \( i, j \in \{1, 2\} \), the system (6.2) has a unique non-negative asymptotically stable steady state \((\lambda^*_1, \lambda^*_2)\) and for this steady state \( \lambda^*_1 > 0 \) and \( \lambda^*_2 > 0 \) that is the unique steady state is positive.

**Proof:**

If either \( \beta_{12} = 0 \) or \( \beta_{21} = 0 \), then it is easy to see that there is a unique steady state and that this steady state is positive. So assume that \( \beta_{12} > 0 \) and \( \beta_{21} > 0 \).
By plugging in (6.5) into (6.4) and rearranging we get the following quadratic equation

$$\beta_{21}\tau_2\lambda_1^2 + (1 + \beta_{12}\tau_1 k_2 - \beta_{21}\tau_2 k_1)\lambda_1 - k_1 = 0 \quad (6.6)$$

Since the product of the leading coefficient and the constant coefficient of the quadratic equation is strictly negative, the quadratic equation has two real roots - one strictly positive and one strictly negative.

Plug in the strictly positive solution $\lambda_1^*$ into (6.5) to get $\lambda_2^* > 0$ which shows that there is a unique non-negative steady state ($\lambda_1^*, \lambda_2^*$) and this steady state is positive.

To study the stability, we consider the Jacobian of (6.2) evaluated at $\Lambda^* = (\lambda_1^*, \lambda_2^*)$ which is

$$J(h(\Lambda^*)) = \begin{pmatrix}
-\frac{1}{\tau_1} - \beta_{12}\lambda_2^* & -\beta_{12}\lambda_1^* \\
-\beta_{21}\lambda_2^* & -\frac{1}{\tau_2} - \beta_{21}\lambda_1^*
\end{pmatrix} \quad (6.7)$$

The determinant of the Jacobian matrix $\det(J)$ is

$$\det(J) = \frac{1}{\tau_1 \tau_2} + \frac{\beta_{12}\lambda_2^*}{\tau_2} + \frac{\beta_{21}\lambda_1^*}{\tau_1} > 0$$

and the trace of the Jacobian $\text{Tr}(J)$ is

$$\text{Tr}(J) = -\frac{1}{\tau_1} - \frac{1}{\tau_2} - \beta_{12}\lambda_2^* - \beta_{21}\lambda_1^* < 0$$

Since $\text{Tr}(J)$ is the sum of the eigenvalues of $J$ and $\det(J)$ is the product of eigenvalues of $J$, it follows that both eigenvalues of $J(h(\Lambda))$ are negative and so from Theorem (6.1.3) it follows that the unique steady state is asymptotically stable. □
**Definition 6.1.6.** For a given value of the firing rates \((\lambda_1, \lambda_2)\) denoted \((\tilde{\lambda}_1, \tilde{\lambda}_2)\), we say that the system is in wake state or that \((\lambda_1, \lambda_2)\) is in wake if \(\tilde{\lambda}_1 > \tilde{\lambda}_2\) and we say that the system is in sleep state or that \((\lambda_1, \lambda_2)\) is in sleep if \(\tilde{\lambda}_1 < \tilde{\lambda}_2\).

For now assume that the wake-active and sleep-active populations have identical time constants \(\tau\) and identical inhibition constants \(\beta\).

**Theorem 6.1.7.** For \(\tau_1 = \tau_2\) and \(\beta_{12} = \beta_{21}\), the unique steady state solution \((\lambda_1^*, \lambda_2^*)\) is in wake if \(k_1 > k_2\) and in sleep if \(k_1 < k_2\).

**Proof:** From equation (6.3), we see that

\[
\frac{\lambda_1 \lambda_2}{\tau_1 \beta_{12}} = \frac{k_1 - \lambda_1}{\tau_1 \beta_{12}} = \frac{k_2 - \lambda_2}{\tau_2 \beta_{21}}
\]

and so using the assumption of equality of time constants and inhibition constants \(\lambda_1 - \lambda_2 = k_1 - k_2\) from which the result follows. \(\square\)

### 6.2 General case of non-negative self-excitation

Now we return to the full two component system (6.1)

\[
\begin{align*}
\lambda_1' &= \frac{k_1 - \lambda_1}{\tau_1} - \beta_{12} \lambda_1 \lambda_2 + \alpha_{11} \lambda_1^2 \left(1 - \frac{\lambda_1}{s_1}\right) \\
\lambda_2' &= \frac{k_2 - \lambda_2}{\tau_2} - \beta_{21} \lambda_2 \lambda_1 + \alpha_{22} \lambda_2^2 \left(1 - \frac{\lambda_2}{s_2}\right)
\end{align*}
\]

where the parameters are \(\alpha_{11}, \alpha_{22} \geq 0, \beta_{12}, \beta_{21} \geq 0, \tau_1, \tau_2 > 0, s_1 > k_1 > 0\) and \(s_2 > k_2 > 0\) and initial conditions are \(0 < \lambda_1(0) < s_1\) and \(0 < \lambda_2(0) < s_2\). When we use a phrase like “for all possible parameter values” we mean for all parameter values and initial conditions that are consistent with these restrictions.
For the system defined above to be a meaningful model of biological significance, the least we require is that the variables $\lambda_1$ and $\lambda_2$ which represent firing rates of action potentials of populations of neurons, do not become negative or conversely do not increase without bound. We will first show that this is indeed the case for the system of differential equations (6.8) for all possible parameter values.

**Theorem 6.2.1.** For all possible parameter values and initial conditions $0 \leq \lambda_1(0) \leq s_1$ and $0 \leq \lambda_2(0) \leq s_2$, if for $t > 0$ $(\tilde{\lambda}_1(t), \tilde{\lambda}_2(t))$ is the unique solution of the system (6.8), then $0 \leq \tilde{\lambda}_1(t) \leq s_1$ and $0 \leq \tilde{\lambda}_2(t) \leq s_2$.

In other words, trajectories are confined to the open rectangle $(0, s_1) \times (0, s_2)$ for $t > 0$.

**Proof:** We only need to show that the flow is inwards on the boundary of the rectangle $(0, s_1) \times (0, s_2)$.

1. $\lambda_1 = 0 \implies \lambda'_1 = \frac{k_1}{\tau_1} > 0$ and so the $x$-component of the flow is to the right on the left edge.

2. $\lambda_2 = 0 \implies \lambda'_2 = \frac{k_2}{\tau_2} > 0$ and so the $y$-component of the flow is upwards on the bottom edge.

3. $\lambda_1 = s_1 \implies \lambda'_1 = \frac{k_1 - s_1}{\tau_1} - \beta_1 s_1 \lambda_2 < 0$ since $s_1 > k_1$ and so the $x$-component of the flow is to the left on the right edge.

4. $\lambda_2 = s_2 \implies \lambda'_2 = \frac{k_2 - s_2}{\tau_2} - \beta_2 s_2 \lambda_1 < 0$ since $s_2 > k_2$ and so the $y$-component of the flow is downwards on the top edge.

This completes the proof. $\square$
Corollary 6.2.2. If there is only one asymptotically stable steady state $x^*$, then the steady state is a global attractor.

Proof: Since all trajectories are confined to the open rectangle $(0, s_1) \times (0, s_2)$ and there are no limit cycles, all trajectories must converge to some stable steady state. Since there is only one stable steady state $x^*$, all trajectories converge to $x^*$, and so $x^*$ is a global attractor. □

Corollary 6.2.3. The unique stable steady state $(\lambda_1^*, \lambda_2^*)$ for zero self-excitation $\alpha_{11} = \alpha_{22} = 0$ is a global attractor.

Proof: Application of the previous corollary. □

The nullclines for system (6.8) are given by

\[
\begin{align*}
\frac{k_1 - \lambda_1}{\tau_1} - \beta_{12} \lambda_1 \lambda_2 + \alpha_{11} \lambda_1^2 \left( 1 - \frac{\lambda_1}{s_1} \right) &= 0 \\
\frac{k_2 - \lambda_2}{\tau_2} - \beta_{21} \lambda_1 \lambda_2 + \alpha_{22} \lambda_2^2 \left( 1 - \frac{\lambda_2}{s_2} \right) &= 0
\end{align*}
\] (6.9)

In fact for non-zero inhibition parameters $\beta_{12}, \beta_{21} \neq 0$, we may solve the first equation in the above system for $\lambda_2$ as a function of $\lambda_1$ to get

\[
\lambda_2 := \phi_1(\lambda_1) = \frac{k_1}{\beta_{12} \tau_1 \lambda_1} - \frac{1}{\beta_{12} \tau_1} + \frac{\alpha_{11}}{\beta_{12}} \lambda_1 \left( 1 - \frac{\lambda_1}{s_1} \right)
\] (6.10)

and similarly for the $\lambda_2$ - nullcline we get

\[
\lambda_1 := \phi_2(\lambda_2) = \frac{k_2}{\beta_{21} \tau_2 \lambda_2} - \frac{1}{\beta_{21} \tau_2} + \frac{\alpha_{22}}{\beta_{21}} \lambda_2 \left( 1 - \frac{\lambda_2}{s_2} \right)
\] (6.11)
6.2.1 Estimates for number of steady state solutions

The following theorem follows directly from the fact that the two component wake-sleep system is a strongly monotone system and the flow on the boundary of the rectangle \((0, s_1) \times (0, s_2)\) is inwards. Since closed orbits and chaotic behavior is ruled out, we must have a fixed point in the interior. We state the theorem and give a direct proof based on intersection of nullclines. Visualising nullclines as graphs of certain functions will be useful in the later results.

**Theorem 6.2.4.** The two-component system (6.8) has at least one steady state solution in the open rectangle \((0, s_1) \times (0, s_2)\) for all possible parameter values \(\alpha_{11} \geq 0, \alpha_{22} \geq 0, s_1 > k_1 > 0, s_2 > k_2 > 0, \beta_{12} \geq 0, \beta_{21} \geq 0, \tau_1 > 0, \tau_2 > 0\).

**Proof:** We assume first that \(\beta_{12}, \beta_{21} > 0\). The statement of the theorem is equivalent to showing that the polynomial system of equations (6.9) has at least one solution consistent with the parameter value restrictions. For the function \(\phi_1(\lambda_1)\) given by (6.10) we note that \(\lambda_2 \to +\infty\) as \(\lambda_1 \to 0^+\). Meanwhile substituting \(\lambda_1 = s_1\) gives \(\lambda_2 = \phi_1(s_1) = \frac{1}{\beta_{12}\tau_1} \left(\frac{k_1}{s_1} - 1\right) < 0\). Thus (6.10) may be viewed as a graph of the function \(\lambda_2 = \phi_1(\lambda_1)\) which when restricted to the domain \((0, s_1)\) has a range containing the interval \([0, \infty)\). Similarly, when the nullcline (6.11) is viewed as a graph of the function \(\lambda_1 = \phi_2(\lambda_2)\) which when restricted to the domain \((0, s_2)\) has a range that contains the interval \([0, \infty)\).

So from straightforward geometrical considerations, we conclude that the two curves defined by \(\phi_1\) and \(\phi_2\) have at least one intersection point inside the open rectangle
\( R := (0, s_1) \times (0, s_2) \) which implies that (6.8) has at least one steady state solution inside the rectangle \( R \) for all possible parameter values.

If \( \beta_{12} = 0 \), then (6.10) is a cubic in \( \lambda_1 \), so the \( \lambda_1 \)-nullclines are just vertical lines, located at a root of the cubic. One of these vertical lines must intersect the \( \lambda_2 \)-nullcline in the closure of the rectangle \( R \), denoted \( \bar{R} \), as can be seen by taking the limit \( \beta_{12} \rightarrow 0 \). However, \( \lambda_1 = 0 \) implies that \( k_1 = 0 \), a contradiction. Also, \( \lambda_1 = s_1 \) implies that \( s_1 = k_1 \), again a contradiction. So, the cubic has a root that is contained in the interval \( (0, s_1) \). This proves that for all possible parameter values there is at least one steady state solution in the open rectangle \( R \). \( \square \)

In fact, it is possible to say a little more than the above theorem. Before we do so, for simplicity’s sake we assume from now on that \( \alpha := \alpha_{11} = \alpha_{22} \), \( \beta := \beta_{12} = \beta_{21} \) and \( \tau_1 = \tau_2 = 1 \), unless otherwise mentioned. Let’s rewrite the dynamical system (6.8)

\[
\begin{align*}
\lambda_1' &= k_1 - \lambda_1 - \beta \lambda_1 \lambda_2 + \alpha \lambda_1^2 \left( 1 - \frac{\lambda_1}{s_1} \right) \\
\lambda_2' &= k_2 - \lambda_2 - \beta \lambda_2 \lambda_1 + \alpha \lambda_2^2 \left( 1 - \frac{\lambda_2}{s_2} \right)
\end{align*}
\]  

(6.12)

For this system, not only does a steady state always exist but also there is at least one \textit{sleep} steady state for \( k_1 < k_2 \) and this steady state is stable.

**Theorem 6.2.5.** For all parameter values \( \{ \alpha \geq 0, \beta \geq 0, s > k_2 > k_1 > 0 \} \), there exists at least one stable steady state solution in sleep.

**Proof:** We construct a bounding region within sleep for trajectories in the phase plane \((c_1, c_2)\). The bounding region we construct is the triangle \( \Delta \) with sides \( \lambda_2 = 0 \),
\[ \lambda_1 = s \text{ and } \lambda_2 = \lambda_1. \] It has already been shown in the proof of theorem (6.2.1) that the flow on the edge \( \lambda_2 = 0 \) is upwards and the flow on the edge \( \lambda_1 = s \) is to the left. Setting \( \lambda_2 = \lambda_1 \) in \( \lambda'_i = h_i(\lambda_1, \lambda_2) \) and then taking the difference of the two equations gives \( \lambda'_1 - \lambda'_2 = k_1 - k_2 < 0 \). This means that the flow on the hypotenuse of the triangle is also inwards. This shows that trajectories with initial condition within the triangle \( \Delta \) never leave the triangle \( \Delta \). Since we know that there are no limit cycles within \( \Delta \), the Poincare-Bendixson theorem implies that there must be an asymptotically stable fixed point within \( \Delta \) which is a sleep steady state. \( \square \)

The following theorem is useful in establishing an upper bound on the number of steady state solutions.

**Theorem 6.2.6.** For \( \beta_{12} > 0 \), the \( \lambda_1 \) - nullcline (6.10) when viewed as a graph of the function \( \lambda_2 = \phi_1(\lambda_1) \) has at most 2 local extrema.

**Proof:** The proof is by ordinary calculus.

\[
\frac{\partial \lambda_2}{\partial \lambda_1} = -\frac{k_1}{\beta_{12} \tau_1 \lambda_1^2} + \frac{\alpha_{11}}{\beta_{12}} \left( 1 - \frac{2\lambda_1}{s_1} \right)
\]

The number of local extrema are just the number of zeros of the expression on the right in the above equation. The zeros in turn are located at the intersection of the two curves \( h_1(\lambda_1) = \frac{k_1}{\tau_1 \lambda_1^2} \) and \( h_2(\lambda_1) = \alpha_{11} \left( 1 - \frac{2\lambda_1}{s_1} \right) \). Since \( h_1 \) is monotonically decreasing and convex while \( h_2 \) is linear and monotonically decreasing, it is clear that \( h_1 \) and \( h_2 \) intersect at exactly 2 points or 0 points (not including the degenerate case where \( h_2 \) is tangent to \( h_1 \) thus resulting in 1 point of intersection). This proves the theorem. \( \square \)
Analogously for $\beta_{21} > 0$, we can argue that the $\lambda_2$ - nullcline viewed as the graph of the function $\lambda_1 = \phi_2(\lambda_2)$ has at most two local extrema. The above theorem sets an upper bound on the number of steady state solutions that the system (6.1) can have. For particular parameter values, if both nullclines do possess the maximum number of local extrema i.e. two local extrema each, a theoretical upper limit on the number of intersections of these curves is 9. Thus there can be at most 9 steady states for the equations (6.1) for any parameter values.

Of course, this theoretical maximum is not necessarily achieved. In fact, we only find at most 5 steady state solutions to (6.1) for any given parameter values. In fact as will be shown, since all bifurcations are saddle-node bifurcations, fixed points appear and disappear in pairs. We have already shown that there exists a steady state solution for any value of $\alpha \geq 0$ in Theorem (6.2.4). We have also shown earlier that the unique steady state for $\alpha = 0$ is stable. It turns out that for any particular parameter values, the number of steady states is 1, 3 or 5 which corresponds to 1, 2 or 3 stable steady states respectively since in general in a saddle node bifurcation, a pair of steady states that appears or disappears has one stable steady state and one unstable steady state.

Recall that for the stochastic system the values of the excitations and inhibitions are restricted to be between 0 and 1 in order for the model to be well-defined i.e. for the rates to remain non-negative. With this restriction on the value of $\alpha$, we can prove a much stronger result for the upper bound on the number of fixed points. We state the main results that will help establish the upper bound.
**Theorem 6.2.7.** For $\beta_{12} > 0$ and $\alpha_{11} < \frac{27k_1}{\tau_1 s_1^2}$,

$$\phi_1(\lambda_1) = \frac{k_1}{\beta_{12} \tau_1 \lambda_1} - \frac{1}{\beta_{12} \tau_1} + \frac{\alpha_{11}}{\beta_{12}} \lambda_1 \left(1 - \frac{\lambda_1}{s_1}\right)$$

is a monotonically decreasing function of $\lambda_1$.

**Proof:** By the previous theorem (6.2.6), $\phi_1$ will be a monotonically decreasing function of $\lambda_1$ if $h_1(\lambda_1) = \frac{k_1}{\tau_1 \lambda_1^2}$ and $h_2(\lambda_1) = \alpha_{11} \left(1 - \frac{2\lambda_1}{s_1}\right)$ have 0 intersections or in other words if $h_1 > h_2$ for all values of $0 \leq \alpha_{11} < \frac{27k_1}{\tau_1 s_1^2}$, since $h_1$ is clearly less than $h_2$ for $\alpha_{11} = 0$. Let $\lambda_1 := x$, $\frac{k_1}{\tau_1} := k$, $\alpha_{11} := \alpha$ and $s_1 := s$. We want to know for what values of $\alpha$ does the equation

$$\frac{k}{x^2} = \alpha \left(1 - \frac{2x}{s}\right)$$

have 0 solutions.

When $\alpha = 0$ the above equation has 0 solutions and for $\alpha$ sufficiently large, it has 2 solutions. So we want to know for what value of $\alpha$ is $y = \alpha \left(1 - \frac{2x}{s}\right)$ a tangent to $y = \frac{k}{x^2}$. Equation of tangent line to $y = \frac{k}{x^2}$ at the point $\mu$ is $y = -\frac{2k}{\mu^3}x + \frac{3k}{\mu^2}$. Setting it equal to $\alpha \left(1 - \frac{2x}{s}\right)$ gives $\alpha = \frac{ks}{\mu^3}$ and $\alpha = \frac{3k}{\mu^2}$ which implies that $\mu = \frac{s}{3}$ and $\alpha = \frac{27k}{s^2}$.

In other words when $\alpha < \frac{27k}{s^2} = \frac{27k_1}{\tau_1 s_1^2}$, $\phi_1$ is a monotonically decreasing function of $\lambda_1$. □

**Corollary 6.2.8.** For $\beta_{21} > 0$ and $\alpha_{22} < \frac{27k_2}{\tau_2 s_2^2}$, $\phi_2(\lambda_2)$ is a monotonically decreasing function of $\lambda_2$. 50
Corollary 6.2.9. For $\alpha_{22} < \frac{27k_2}{\tau_2 s_2^2}$, $\phi_2(\lambda_2)$ is invertible and $\lambda_2 = \phi_2^{-1}(\lambda_1)$ is a monotonically decreasing function.

The parameter values used in numerical simulation were $k_1 = 5$, $k_2 = 5.5$, $\tau_1 = \tau_2 = 1$ and $s_1 = s_2 = 10$, so for instance $\frac{27k_1}{\tau_1 s_1^2} = 1.35$. Since $0 \leq \alpha_{11} \leq 1$, the condition of the theorem is satisfied and $\phi_1$ is a decreasing function and so is $\phi_2$.

Theorem 6.2.10. $\lambda_j = \phi_i(\lambda_i)$ for $i, j \in \{1, 2\}, i \neq j$ has at most one inflection point in the interval $[0, s_i]$.

Proof:

$$\frac{\partial^2 \lambda_j}{\partial \lambda_i^2} = \frac{2k_i}{\beta_{ij} \tau_i \lambda_i^3} - \frac{2\alpha_{ii}}{\beta_{ij} s_i} = 0$$

has only one solution $\lambda_i^* = \left(\frac{k_i s_i}{\tau_i \alpha_{ii}}\right)^{\frac{1}{3}}$ and $\phi_i$ is convex on $[0, \lambda_i^*]$ and concave on $[\lambda_i^*, s_i]$. □

Corollary 6.2.11. $\lambda_i = \phi_i^{-1}(\lambda_j)$ has at most one inflection point in the interval $[0, s_j]$ and $\phi_i^{-1}$ is concave to the left of the inflection point and convex to the right.

For the range of parameters used in the model, we find that (6.8) can only have 1 stable steady state or 3 steady states of which 2 are stable and 1 is unstable. This will be clear in the bifurcation diagrams in a later section. In the next section, we show that the 1 stable steady state is necessarily in sleep and after the bifurcation point, the two new steady states that emerge are necessarily in wake.
6.2.2 Wake-sleep classification for steady state solutions

**Theorem 6.2.12.** For $\alpha := \alpha_{11} = \alpha_{22}$, $\beta := \beta_{12} = \beta_{21}$, $\tau_1 = \tau_2 = 1$ and $s := s_1 = s_2$, if $k_1 \neq k_2$, then the steady state solutions of the two component system (6.12) are either in wake or in sleep.

*Proof:* The nullclines of (6.12) with $s = s_1 = s_2$ are

\[
\begin{align*}
    k_1 - \lambda_1 - \beta \lambda_1 \lambda_2 + \alpha \lambda_1^2 \left(1 - \frac{\lambda_1}{s}\right) &= 0 \\
    k_2 - \lambda_2 - \beta \lambda_1 \lambda_2 + \alpha \lambda_2^2 \left(1 - \frac{\lambda_2}{s}\right) &= 0
\end{align*}
\] (6.13)

it is easy to see that if $\lambda_1^* = \lambda_2^*$ then $k_1 = k_2$. So if $k_1 \neq k_2$, then either $\lambda_1^* > \lambda_2^*$ or $\lambda_1^* < \lambda_2^*$ which shows that the steady states are either in wake or in sleep. □

The solutions $(\lambda_1, \lambda_2)$ of the polynomial system of equations (6.13) depend continuously on the parameters $\alpha, \beta, s, k_1$ and $k_2$. We have shown previously that there exists a unique solution for the above system which is in sleep state for $k_1 < k_2$ and $\alpha = 0$. As $\alpha$ is increased gradually, we can follow this steady state solution as a function of $\alpha$. The previous theorem tells us that this particular steady state solution which was in sleep for $\alpha = 0$ continues to be in sleep for all positive values of $\alpha$ since it is not allowed to cross the diagonal $\lambda_1 = \lambda_2$. Of course, a bifurcation may occur which may result in loss of the steady state as $\alpha$ increases. However, if two steady states are on the same bifurcation curve with respect to any bifurcation parameter, both the steady states are either in wake or both are in sleep. We prove a more general statement in the following definitions and theorem.
Set of all $D$ fixed points

Consider a dynamical system $\Lambda' = h(\Lambda, A)$ with variables $\Lambda := (\lambda_1, \lambda_2, \ldots, \lambda_n) \in \mathbb{R}^n$ and parameters $A := (\alpha_1, \alpha_2, \ldots, \alpha_m) \in \mathbb{R}^m$. We define the set-valued function that maps each point in the parameter space to a set of fixed points of the differential equations.

$$\rho : A \mapsto \{\Lambda^*| h(\Lambda^*, A) = 0\}$$

We assume that the domain of $\rho$ is a closed box $D \subset \mathbb{R}^m$.

**Definition 6.2.13.** For a domain $D \subset \mathbb{R}^m$, define the range of $\rho$ to be the set of all $D$ fixed points, denoted by $M_{\rho(D)}$, or simply $M_D$. When the domain $D$ is simply the domain of one parameter $\alpha$, we will denote the set of all $D$ fixed points by $M_\alpha$ and refer to it as the set of all $\alpha$ fixed points.

**Definition 6.2.14.** Let $M \subset \mathbb{R}^n$ and let $x, y \in M$. If there exists a continuous path $P_{xy}$ from $x$ to $y$ such that $P_{xy} \subset M$, then we say that $x$ and $y$ are on the same connected component of $M$.

*Note:* A system of differential equations has the property that all solutions, including steady state solutions, depend continuously on the parameters. This ensures that in most cases the set of all fixed points has non-trivial topology. By trivial topology, we mean discrete topology.

We now return to the wake-sleep system. If the domain of $\rho$ is the domain of a single bifurcation parameter say the self-excitation parameter $\alpha$, a connected component
of the set of all $\alpha$ fixed points of the wake-sleep system (6.12) may be thought of as a parametrized curve in the phase plane $(\lambda_1, \lambda_2)$ parametrized by $\alpha$. When the domain includes more than 1 bifurcation parameter, the set of fixed points and its each connected component will have a more complicated topology. In any case, the following theorem is a very useful tool in understanding the qualitative behavior of the system.

**Theorem 6.2.15.** For the system (6.12) with $s = s_1 = s_2$, the property of being in wake or being in sleep is a topological property of the set of all $D$ fixed points.

By this statement, we mean that all the fixed points on one connected component of the set of all $D$ fixed points lie either in wake or in sleep.

**Proof:** Consider the set of all $D$ fixed points $M_D$, which has a connected component $M_{D_1}$ with more than one point. If no such component exists, there is nothing to prove. Suppose at least one point $x$ on the connected component $M_{D_1}$ is in wake and at least one point $y$ is in sleep. Then there exists a continuous path $P_{xy}$ connecting the two points such that $P_{xy} \subset M_{D_1}$. By the intermediate value theorem, there exists a point $z \in M_{D_1}$ such that $\lambda_1(z) = \lambda_2(z)$ which contradicts Theorem (6.2.12). So all fixed points on $M_{D_1}$ must be either in wake or must be all in sleep. $\square$

Let $M_\alpha$ be the set of all $\alpha$ fixed points for the parameter $\alpha$. Define the projection operator $P_{\lambda_i}(M_\alpha)$ as the projection onto the component $\lambda_i$ of the set of all $\alpha$ fixed points. The ‘graph’ of the set-valued function $\alpha \mapsto \alpha \times P_{\lambda_i}(M_\alpha)$ is called a bifurcation diagram for the parameter $\alpha$ and the variable $\lambda_i$. Such a diagram is represented for instance in (6.2) and (6.3).
An important consequence of the above theorem is that when we follow a wake fixed point along a bifurcation curve, it cannot change to a sleep fixed point. In particular, if two fixed points are created or annihilated both must be wake or both must be sleep.

**Significance of wake-sleep classification of states**

When we simulate the original stochastic dynamical system (3.1), we find the system spends a significant amount of time near the steady states of the deterministic system. These firing rates are often referred to as *meta-stable states* in the stochastic setting, although it should be pointed out that the meta-stable states are not steady states in the same sense as in a deterministic dynamical system since the probability of a sample path converging to a fixed point for all future time is zero.

The significance of the results from bifurcation analysis of the deterministic system to the stochastic dynamical system (3.1) is that the stochastic system spends most of its time near the fixed points of the deterministic system. For a large range of self-excitation parameter $\alpha$, near these fixed points, either wake firing is significantly higher than sleep firing or sleep firing is significantly higher than wake firing. Biologically, this corresponds to the two states being behaviorally distinguishable in a mammal, and little overlap between the two states. In fact, we will see situations where there may be difficulties in distinguishing the two states, and these could correspond to sleep-related pathological conditions or dyssomnias such as narcolepsy. Narcolepsy is a condition where behavior normally associated with sleep, in particular REM sleep, such as muscle atonia invades the wake state causing temporary paralysis or cataplexy. The model is not complex enough to predict particular pathological
outcomes or symptoms, however disregulation in the features that are present in the model such as mutual inhibition and self-excitation can suggest possible underlying mechanisms for a class of dyssomnias.

6.3 Bifurcation analysis of steady state solutions for two component model

6.3.1 Bifurcation parameter is self-excitation $\alpha = \alpha_{11} = \alpha_{22}$

We demonstrate the existence of the steady state solutions for the two component (sleep active and wake active) model. We define the self-excitation parameter $\alpha := \alpha_{11} = \alpha_{22}$ and make it the bifurcation parameter. The dependent variables are the steady state firing rates of the wake-active population $\lambda_1$ and the sleep-active population $\lambda_2$. We fix the inhibition parameter $\beta = \beta_{12} = \beta_{21} = 0.5$, $k_1 = 5$ and $k_2 = 5.5$. For the self-excitation parameter $\alpha = 0, 0.5, 1.6, 1.8, 2.0$ we find that the nullclines have 1, 3, 5, 3, 1 intersections respectively.

As we vary the parameter $\alpha$, at a bifurcation point either a pair of steady states of which one is stable and one is unstable either appears or disappears. Precisely at the bifurcation point we have the degenerate case where the two steady states ‘collide’ and there is only one semi-stable steady state. If we exclude this point, then there can be only 1, 3 or 5 steady states. The behavior of these can be studied more easily using a bifurcation diagram.

We now describe the main ‘phases’ of the bifurcation diagram with respect to $\alpha$ as it
increases. The following comments should be read while referring to the bifurcation diagrams (6.2) and (6.3).

1. As observed earlier, for small values of self-excitation $\alpha$, there is only one steady state which is stable. The location of this steady state is in sleep for $k_1 < k_2$ (here $k_1 = 5$ and $k_2 = 5.5$). As $\alpha$ increases, firing rate $\lambda_2$ of the sleep-active population increases significantly and meanwhile the firing rate $\lambda_1$ of the wake-active population at the steady state decreases. More significantly, as can be seen from (6.2) and (6.3) at approximately $\alpha = 0.37$ a saddle-node bifurcation occurs and we get a new pair of steady states, one stable and one unstable. Both the new steady states are wake states.

We note that after the bifurcation occurs, there is precisely one stable steady state that is in sleep (the upper branch for $\lambda_2$ and the lower branch for $\lambda_1$) and one stable steady state that is in wake (upper branch for $\lambda_1$ and lower branch for $\lambda_2$).

2. For approximately $\alpha = 1.5$ another bifurcation occurs in the original sleep state and so we get two stable steady states in sleep and one unstable steady state in sleep. This along with the two wake states makes a total of five steady states of which 3 are stable.

3. A third bifurcation occurs at approximately $\alpha = 1.78$. The two wake states collide and disappear, and so the only stable steady states remaining, two of them, are both in sleep. Although the lower firing rate for $\lambda_2$ corresponds to
sleep, we note that $\lambda_1$ is also firing at a relatively high rate at this steady state. Such a disregulation may be thought of as resulting from ‘excessive’ self-excitation.

4. A fourth bifurcation occurs at approximately $\alpha = 1.95$. A stable and an unstable sleep steady state collide and annihilate each other. So we are only left with one sleep steady state with both wake-active and sleep-active populations firing at relatively high rates.

An important observation to make is the robustness in the system response to large variations in parameters. Robustness is a very desirable feature in biological systems. Biological systems, whether they are individuals animals, individuals cells, populations of cells, organs or some other entity, often tend to have large variations in parameters. Not only that, the parameters themselves may be changing as the animal develops. Thus what may be considered a parameter for an animal at a particular age or a particular time of the day, may really be slowly-changing variables that depend on the age of the animal or circadian effects. A system that has dramatically different behavior for a slight change in parameters is likely to have low fitness and hence is likely not to be preserved over an evolutionary time scale.

For the particular model in hand, we note that for a large range of the self-excitation parameter i.e. for $0.6 < \alpha < 1.6$, we have two stable steady states, one in sleep and one in wake. Not only that, but the two steady states have firing rates that do not vary too much with $\alpha$. This indicates that changing $\alpha$ within this range will not
have a large effect on behavior of the individual. This is precisely the property of robustness which we desire our model to possess.

An interesting phenomenon arises when the self-excitation parameter exceeds this range. We get a new pathological sleep steady state which has high firing rates for both the wake-active and the sleep-active populations. A natural candidate for a biological interpretation would be a state of dyssomnia due to excessive self-excitation.

Having said that, we should sound a note of caution that we haven’t demonstrated the validity of the model for large values of self-excitation, or whether the deterministic model continues to be a reasonable approximation for large values of jumps.

6.3.2 Bifurcation parameter is mutual inhibition $\beta = \beta_{12} = \beta_{21}$

In figures (6.4) and (6.5) we hold the self-excitation parameter $\alpha$ constant at the values 0, 0.5 and 1.6 and draw a bifurcation diagram of fixed points for the bifurcation parameter $\beta$.

1. For $\alpha = 0$ there is a unique steady state which is stable and in sleep. Both firing rates for this steady state decrease as $\beta$ increases.

2. For non-zero self-excitation $\alpha$, we get more interesting behavior. For $\alpha = 0.5$ the firing rate for $\beta = 0$ is very close to the maximum firing rate. But it drops rapidly to close to zero when $\beta$ is increased from 0 to 0.5. Around $\beta = 0.35$, a bifurcation occurs and one stable fixed point and one unstable fixed point appear, both the fixed points are in wake.
When we follow the sleep firing rate $\lambda_2$ as $\beta$ increases we observe that first the stable steady state $\lambda_2$ decreases but then it increases again. This may seem surprising at first. However it makes sense when we consider that increasing $\beta$ not only inhibits the sleep-active population but also disinhibits it by inhibiting the wake-active population sending inhibition to the sleep-active population.

3. For $\alpha = 1.6$, we find a brief range of $\beta$ for which 5 steady states co-exist of which 3 are stable.

The main reason for studying the bifurcation diagram with respect to $\beta$ is to demonstrate robustness with respect to $\beta$. For instance, for $\alpha = 0.5$, for any value of $\beta > 0.5$, we have 3 steady states of which 2 are stable, one in wake and one in sleep. The firing rate of the two populations does not vary much with large variations in $\beta$. Especially notice that the firing rate of the active population (DLPT during wake state and PnO during sleep state) is almost constant for large variations in $\beta$. As has been mentioned earlier, the property of robustness is highly desirable in biological modeling to account for individual variations.

6.3.3 Bifurcation diagrams for the case when $\alpha_{11} \neq \alpha_{22}$

In figures (6.6) and (6.7) we fix the mutual inhibition parameter $\beta = 0.5$ and sleep self-excitation is chosen to be $\alpha_{22} = 0.5$ and $\alpha_{22} = 1.6$ while $\alpha_{11}$ is the bifurcation parameter.

When the two firing rates $\alpha_{11}$ and $\alpha_{22}$ are close to each other, we find that there are two distinct stable steady states, one in sleep and one in wake. However, when
the self-excitation of wake is significantly smaller than that of sleep say, $0 \leq \alpha_{11} < 0.3, \alpha_{22} = 0.5$, there is only stable steady state which is in sleep. Similarly when wake self-excitation is relatively large $\alpha_{11} > 1.2, \alpha_{22} = 0.5$, then there is only one stable steady state which is in wake. Thus, as one would expect, if one of the two populations has excessive self-excitation relative to the other population, even in the stochastic setting the system may not be able to switch between states and may remain in wake or in sleep for most of the time.

Note: The labels $c_i$ in the graphs are synonymous to the firing rate variables $\lambda_i$ and the parameters $exc$ and $inh$ are synonymous with $\beta$ and $\alpha$ respectively.
Figure 6.1: Nullclines in phase plane ($\lambda_1, \lambda_2$) for mutual inhibition parameter $\beta = 0.5$.

The values of self-excitation parameter $\alpha$ are $0, 0.5, 1.6, 1.8$ and $2.0$ respectively.
Figure 6.2: Wake firing rate (\(\lambda_1\)) steady states as a function of self-excitation parameter (\(\alpha\)) for \(\beta = 0.5\)

Figure 6.3: Sleep firing rate (\(\lambda_2\)) steady states as a function of self-excitation parameter (\(\alpha\)) for \(\beta = 0.5\)
Figure 6.4: Wake firing rate ($\lambda_1$) steady states as a function of mutual inhibition parameter ($\beta$) for $\alpha = 0, 0.5, 1.6$

Figure 6.5: Sleep firing rate ($\lambda_2$) steady states as a function of mutual inhibition parameter ($\beta$) for $\alpha = 0, 0.5, 1.6$
Figure 6.6: Wake firing rate ($\lambda_1$) steady states as a function of wake self-excitation ($\alpha_{11}$) for $\beta = 0.5$ and $\alpha_{22} = 0.5, 1.6$

Figure 6.7: Sleep firing rate ($\lambda_2$) steady states as a function of wake self-excitation ($\alpha_{11}$) for $\beta = 0.5$ and $\alpha_{22} = 0.5, 1.6$
CHAPTER 7
SIMULATION RESULTS OF THE STOCHASTIC MODEL
FOR THE TWO COMPONENT SYSTEM

7.1 Zero Self-excitation

We numerically computed the solution of the stochastic system (3.8) for the inhibition parameter \(\beta = 0.5\) and self-excitation \(\alpha = 0\). Figure (7.1) is a portion of the sample path taken by the firing rates \(\lambda_1\) and \(\lambda_2\). Both the firing rates may be thought of as small changes around their equilibrium values. However, these small changes should not be thought of as a random walk, since besides a restoring force towards an autonomous firing rate, there are jumps resulting from inhibition. However, the magnitude of the jumps is small precisely near the equilibrium values and so we expect the stochastic system to have a mode near the value of the stable steady state in the deterministic system.

The modes can be seen more clearly in the histogram plots of the firing rates. We ran a simulation for 20,000 seconds of model time. The \(x\)-axis in (7.2) represents the firing rate of the two populations. The unimodal distribution of the firing rates roughly corresponds to the fact that the deterministic model has one stable steady
state and we expect the location of the mode in the stochastic model to be close to the value predicted by the steady state value in the deterministic model. 

The results of the simulation may be summarized using different averages such as mean, median and mode, which are presented in the following table (7.1). We have also presented the value of the steady state of the deterministic model for comparison. 

A portion of the phase plane trajectory in the simulation for $t \in [100, 1000]$ is shown in (7.3). We notice in the picture that although the firing rates $\lambda_1$ and $\lambda_2$ appear to be negatively correlated, as we expect for mutual inhibition, the correlation appears to be quite weak.

Figure 7.1: Time course of $\lambda_1$ (in black) and $\lambda_2$ (in red); $x$-axis: time in seconds, $y$-axis: firing rate per second
Figure 7.2: Histograms for firing rates of the wake-active population ($\lambda_1$) and the sleep active population ($\lambda_2$) for $\alpha = 0$

To quantify the previous remark, we calculate the correlation matrix for the vectors of firing rates $\lambda_1(t)$ and $\lambda_2(t)$ for the parameter values $\beta = 0.5$ and $\alpha = 0$.

$$
\begin{pmatrix}
1.0550 & -0.6363 \\
-0.6363 & 1.4077
\end{pmatrix}
$$

7.2 Positive Self-excitation

We carry out analysis similar to the previous section, now for positive self-excitation. We choose the parameters $\beta = 0.5$ and self-excitation $\alpha = 0.5$. A portion of the time course of the firing rates in the simulation is presented in the figure (7.4). We observe here that the presence of distinct sleep and wake states is more obvious even at a
Table 7.1: Average values of the firing rates in the simulation for $\beta = 0.5$ and $\alpha = 0$

<table>
<thead>
<tr>
<th></th>
<th>$\lambda_1$</th>
<th>$\lambda_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>2.2266</td>
<td>2.7573</td>
</tr>
<tr>
<td>Median</td>
<td>2.0804</td>
<td>2.6288</td>
</tr>
<tr>
<td>Mode</td>
<td>1.62</td>
<td>2.222</td>
</tr>
<tr>
<td>Fixed point</td>
<td>2.1504</td>
<td>2.6504</td>
</tr>
</tbody>
</table>

glance, compared to the case of zero self-excitation. Having a positive self-excitation drives the two firing rates, those of wake-active and sleep-active populations further apart. This results in two distinct stable steady states, one in wake and one in sleep, in the deterministic model. In the stochastic model, it almost eliminates the possibility of the two populations firing at comparable firing rates for a significant period of time, thus making the sleep and wake states clearly distinguishable.

In the figures (7.5) is the histogram plot of the firing rates of the wake-active and sleep-active populations respectively. Unlike in the case of zero self-excitation, we find that the distribution of firing rates is not unimodal. In fact, we find two large peaks near opposite ends of the spectrum and we also find 3 smaller peaks in between. The two tallest peaks roughly correspond to the two stable steady states found in the deterministic model for $\alpha = 0.5$. It is interesting to find the three smaller peaks in between. When we examine the vector field in the deterministic case, we find that
between the two stable steady states, there are 3 locations where the magnitude of the velocity vectors becomes near zero. One may speculate that this is the reason for the small peaks in the stochastic case. So the smaller peaks may be also thought of ‘meta-stable’ states of the deterministic model. Meta-stable states, even though they are not proper fixed points, are ‘bottlenecks’ where the system can spend a significant portion of time if it reaches there.

Even though the bistability of the deterministic model translates into bimodality of the stochastic model, the location of the modes is not too close to the stable fixed points in the deterministic model. This is not too surprising because the magnitude of jumps in the firing rates due to incoming spikes is so large that the deterministic approximation is not strictly valid. However, it is interesting to observe the qualitative map of bistability predicting bimodality in the stochastic model.
Figure 7.4: Time course of $\lambda_1$ (in black) and $\lambda_2$ (in red); $x$-axis: time in seconds, $y$-axis: firing rate per second

A portion of the phase plane trajectory in the simulation for $t \in [100, 1000]$ is shown in (7.6). By examining the diagram, we notice now that the negative correlation between $\lambda_1$ and $\lambda_2$ is much larger in magnitude.

For $\beta = 0.5$ and $\alpha = 0.5$, we calculate the correlation matrix for the vectors of firing rates $\lambda_1(t)$ and $\lambda_2(t)$ to be

$$
\begin{pmatrix}
9.8512 & -8.9917 \\
-8.9917 & 11.1352
\end{pmatrix}
$$

As we would expect, there is a much stronger self-correlation of both $\lambda_1$ and $\lambda_2$ and a much stronger negative correlation between the two.
Table 7.2: Average firing rates in simulation for $\beta = 0.5$ and $\alpha = 0.5$

<table>
<thead>
<tr>
<th></th>
<th>$\lambda_1$</th>
<th>$\lambda_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>3.5905</td>
<td>4.7192</td>
</tr>
<tr>
<td>Median</td>
<td>2.1870</td>
<td>4.1092</td>
</tr>
<tr>
<td>Largest Modes</td>
<td>0.6453</td>
<td>9.355</td>
</tr>
<tr>
<td></td>
<td>9.207</td>
<td>0.8334</td>
</tr>
<tr>
<td>Smaller modes</td>
<td>4.56</td>
<td>6.098</td>
</tr>
<tr>
<td></td>
<td>5.699</td>
<td>5.847</td>
</tr>
<tr>
<td></td>
<td>6.769</td>
<td>4.678</td>
</tr>
</tbody>
</table>

Table 7.3: Location of fixed points in the deterministic model

<table>
<thead>
<tr>
<th></th>
<th>$\lambda_1$</th>
<th>$\lambda_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stable fixed points for $\beta = 0.5$ and $\alpha = 0.5$</td>
<td>1.1401</td>
<td>7.7813</td>
</tr>
<tr>
<td></td>
<td>7.2733</td>
<td>1.3581</td>
</tr>
<tr>
<td>Unstable fixed point</td>
<td>3.763</td>
<td>3.0044</td>
</tr>
</tbody>
</table>
Figure 7.5: Histograms for firing rates of the wake-active population and the sleep active population for $\alpha = 0.5$

Figure 7.6: Phase plane $(\lambda_1, \lambda_2)$ for $\alpha = 0.5$
CHAPTER 8
ANALYSIS OF THE THREE COMPONENT
DETERMINISTIC SYSTEM

Earlier we studied the two component wake-sleep system, where we showed that there exist a range of parameters for which there are two stable steady states, one in wake and one in sleep. The wake and sleep state occupancy times were both exponentially distributed with the only difference being the mean occupancy time which is the unique parameter in an exponential distribution. The chapter on the two component wake-sleep system may be viewed as a study of mutual inhibition.

Now we wish to study mutual excitation between the wake-active and wake-promoting populations before integrating this understanding with that coming from mutual inhibition to get a complete picture of the three component wake-sleep system.

8.1 Two component mutually excitatory system

In this section we are interested in the mutually excitatory part of the circuit composed of the wake-active and the wake-promoting populations. The sleep-active population inhibits both the wake-active and the wake-promoting populations. For the purposes of this section, to understand the phenomenon of mutual excitation, we
make the sleep firing rate $\lambda_2$ into a bifurcation parameter. We will also assume that the self-excitation parameters $\alpha_{11} = \alpha_{22} = 0$.

The deterministic system reduces to

$$
\begin{align*}
\lambda'_1 &= k_1 - \lambda_1 - \beta_{12}\lambda_1\lambda_2 + \alpha_{13}\lambda_1 \left( 1 - \frac{\lambda_1}{s_1} \right) \lambda_3 \\
\lambda'_3 &= k_3 - \lambda_3 - \beta_{32}\lambda_3\lambda_2 + \alpha_{31}\lambda_3 \left( 1 - \frac{\lambda_3}{s_3} \right) \lambda_1
\end{align*}
$$

(8.1)

Let $1 + \beta_{12}\lambda_2 = 1 + \beta_{32}\lambda_2 = \mu \in [1, 1 + s_2]$ be the bifurcation parameter and so

$$
\begin{align*}
\lambda'_1 &= k_1 - \mu\lambda_1 + \alpha_{13}\lambda_1 \left( 1 - \frac{\lambda_1}{s_1} \right) \lambda_3 \\
\lambda'_3 &= k_3 - \mu\lambda_3 + \alpha_{31}\lambda_3 \left( 1 - \frac{\lambda_3}{s_3} \right) \lambda_1
\end{align*}
$$

(8.2)

It is easy to show that there exists at least one stable steady state in the interior of $[0, s_1] \times [0, s_3]$. We first show that the flow is inwards.

**Theorem 8.1.1.** The flow on the boundary of $[0, s_1] \times [0, s_3]$ is inwards.

*Proof:* When $\lambda_1 = 0$, $\lambda'_1 = k_1 > 0$. $\lambda_1 = s_1$ implies that $\lambda'_1 = k_1 - \mu s_1 \leq k_1 - s_1 < 0$ because $\mu \geq 1$. A similar calculation for $\lambda_3$ proves the theorem. □

The mutually excitatory system is monotone and therefore does not have limit cycles. This fact along with the above theorem show that there must be a stable steady state in the interior of $[0, s_1] \times [0, s_3]$.

We observe in (8.1) and (8.2) the effect of changing the sleep-firing rate. Suppose that the sleep-firing rate is high. For small changes in $\lambda_2$, there is very little change
Figure 8.1: Wake firing rate ($\lambda_1$) at steady state as a function of sleep firing rate ($\lambda_2$) for $\alpha_{13} = \alpha_{31} = 0.8$ and $\alpha_{11} = \alpha_{22} = 0$, $\beta_{12} = \beta_{32} = 0.5$

in $\lambda_1$ and especially in $\lambda_3$. Thus the system remains in sleep for small changes in $\lambda_2$. However if a large fluctuation occurs and $\lambda_2$ is close to 0, a state change from sleep to wake can occur and $\lambda_1$ and $\lambda_3$ will suddenly jump to a higher firing rate.

*Note:* The labels $c_i$ in the graphs are synonymous to the firing rate variables $\lambda_i$ and the parameters $exc$ and $inh$ are synonymous with $\beta$ and $\alpha$ respectively.

### 8.2 Effect of external excitation on two component wake-sleep system

As the animal develops, the third important component of the sleep-wake circuit, namely the wake-promoting population of Locus Coeruleus comes into play. In order
to understand the effect of LC on the wake-sleep system, we make the firing rate $\lambda_3$ into a parameter and study what happens to the basic sleep-wake switch as $\lambda_3$ is increased.

$$\lambda'_1 = \frac{k_1 - \lambda_1}{\tau_1} - \beta_1 \lambda_1 \lambda_2 + \alpha_{11} \lambda_1^2 \left(1 - \frac{\lambda_1}{s_1}\right) + \alpha_{13} \lambda_1 \left(1 - \frac{\lambda_1}{s_1}\right) \lambda_3$$

$$\lambda'_2 = \frac{k_2 - \lambda_2}{\tau_2} - \beta_2 \lambda_1 \lambda_2 + \alpha_{22} \lambda_2^2 \left(1 - \frac{\lambda_2}{s_2}\right)$$

(8.3)

### 8.2.1 Zero self-excitation

Let $\tau_1 = \tau_2 = 1$, $\alpha_m := \alpha_{13} \lambda_3 \in [0, s_3]$ is the external excitation parameter. Also for now let $\alpha_{11} = \alpha_{22} = 0$ i.e. zero self-excitation and so the above system becomes
\[
\lambda_1' = k_1 - \lambda_1 - \beta_{12}\lambda_1\lambda_2 + \alpha_m\lambda_1 \left(1 - \frac{\lambda_1}{s_1}\right)
\]
\[
\lambda_2' = k_2 - \lambda_2 - \beta_{21}\lambda_1\lambda_2
\]

(8.4)

The nullclines of this system can be solved for \(\lambda_2\).

\[
\lambda_2 = \frac{1}{\beta} \left(\frac{k_1}{\lambda_1} - \frac{\alpha_m}{s_1}\lambda_1 + \alpha_m - 1\right) := f_1(\lambda_1)
\]
\[
\lambda_2 = \frac{k_2}{1 + \beta\lambda_1} := f_2(\lambda_2)
\]

**Theorem 8.2.1.** The system (8.4) has a unique globally asymptotically stable steady state for all parameter values.

**Proof:** It is easy to see that both \(f_1\) and \(f_2\) are decreasing functions of \(\lambda_1\) and so have at most one intersection point. Also observe that \(f_2(0) = k_2 \in [0, s_2]\) and \(f_2(s_1) \in [0, k_2] \subset [0, s_2]\). On the other hand, \(f_1(0^+) = \infty\) and \(f_1(s_1) = \frac{1}{\beta} \left(\frac{k_1}{s_1} - 1\right) < 0\). So \(f_1\) and \(f_2\) must have at least one intersection point.

Using arguments similar to the ones used in proof of (6.2.1) we can show that the flow is inwards and so the unique steady state is globally asymptotically stable. □

We have shown that for each \(\alpha_m\), there is a unique \((\lambda_1^*, \lambda_2^*)\) such that \(\lambda_1' = \lambda_2' = 0\) at that point. Define the function \(x = \lambda_1^* = h(\alpha_m)\)

**Theorem 8.2.2.** \(h\) is an increasing function of \(\alpha_m\).
Proof: Set $f_1 = f_2$ and it is easy to show by implicit differentiation that $\frac{dh}{d\alpha_m} = \frac{1 - \lambda_1}{s_1 + \frac{k_1}{\lambda_1}}$ which is clearly non-negative for $\lambda_1 \in [0, s_1]$. □

As $\alpha_m$ increases the $\lambda_2$ nullcline remains unchanged and so the fixed point slides down the $\lambda_2$ nullcline. Referring to figure (8.3) we note that as $\alpha_m = \alpha_{13} \lambda_3$ is increased from minimum of 0 to maximum of $s_3$, the wake firing rate $\lambda_1$ increases from approximately 2.1 to 9.1. This shifts the unique steady state into wake. Correspondingly, the sleep firing rate $\lambda_2$ also decreases with increase in $\alpha_m$.

Figure 8.3: Wake firing rate ($\lambda_1$) at the unique steady state as a function of wake-promoting firing rate ($\lambda_3$) for $\beta = 0.5$ and $\alpha_{11} = \alpha_{22} = 0, \alpha_{13} = 0.8$
Figure 8.4: Sleep firing rate ($\lambda_2$) at the unique steady state as a function of wake-promoting firing rate ($\lambda_3$) for $\beta = 0.5$ and $\alpha_{11} = \alpha_{22} = 0, \alpha_{13} = 0.8$

### 8.2.2 Positive self-excitation

Now we consider the case of $\alpha_{11} > 0, \alpha_{22} > 0$.

**Theorem 8.2.3.** The system (8.3) has at least one stable steady state in $[0, s_1] \times [0, s_2]$ for all values of $\alpha_{11} \geq 0, \alpha_{22} \geq 0$.

*Proof:* The proof is similar to (6.2.1), we show that the flow is inwards at the boundaries. □

In fact in the bifurcation diagrams (8.5) and (8.6), we notice that this case is qualitatively different from the case of zero self-excitation. For small values of the bifurcation parameter $\lambda_3$ there is bistability, however for large values there is only one stable steady state in wake. For $\lambda_3$ to increase, the three component system must
already be in wake state, which means that the $\lambda_1$ firing rate is going to be in the upper branch (higher $\lambda_1$). The firing rate $\lambda_1$ increases from 7.1 to 9.1 as external excitation is increased, but this increase is much smaller compared to the case of zero self-excitation.

Increase in wake firing rate while the system is in wake is how the wake bouts acquire a power-law like distribution. Even though the bifurcation diagram for the non-zero self-excitation looks qualitatively different from that of zero self-excitation, the emergence of power law is due to the same reason in each. The firing rate of the wake-active population increases when the system is in wake. Since this is the main effect that we want to study, we will for simplicity assume from here on that $\alpha_{11} = \alpha_{22} = 0$.

Figure 8.5: Wake firing rate ($\lambda_1$) at the unique steady state as a function of wake-promoting firing rate ($\lambda_3$) for $\beta = 0.5$ and $\alpha_{11} = \alpha_{22} = 0.5, \alpha_{13} = 0.8$
8.3 Bifurcation parameter is mutual excitation parameter -
Effect of development on wake distribution

In this section, the mutual excitation parameters $\alpha_{13}$ and $\alpha_{31}$ will be treated as bifurcation parameters. The bifurcation diagrams (8.7), (8.8) and (8.9) have $\alpha_{13}$ as the bifurcation parameter while in (8.10), (8.11) and (8.12) we treat $\alpha := \alpha_{13} = \alpha_{31}$ as the bifurcation parameter. The maximum firing rate of wake $s_1 = 20$ and that of sleep and LC is $s_2 = s_3 = 10$.

Particularly important is the bifurcation diagram with respect to the mutual excitation parameter $\alpha$ since increasing $\alpha$ represents age related development. Even though
in the dynamical system (4.24), $\alpha$ is a fixed parameter, the system (4.24) holds on the timescale of minutes or hours. Over a developmental time scale of days, we expect the coupling constants $\alpha_{ij}$ and $\beta_{ij}$ to change. This change in parameters is slow compared to the change in other variables and so can be treated as constants for most purposes. However if we want to understand the changes happening through development, we need to interpret the $\alpha$ variable in (8.10), (8.11) and (8.12) as an age variable.

If parameters are not true parameters but rather slow variables, the obvious question to ask is: “What are the dynamics of the slow variables?”. The answer to this question can be given in different ways. One way would be to acquire a detailed understanding of the biological phenomena that are responsible for the dynamics of the slow variables. The change in the coupling coefficients is a function of the developmental age of the animal. As the animal develops several biophysical changes are occurring simultaneously. These changes include but are not limited to, increase in number of neurons in a population, increasing strength of connection between neurons within a population, increasing strength of connection among certain target populations. A biophysically realistic model that includes all these changes is beyond the scope of our study, also most details about the biological system are not available and are generally difficult to obtain due to large individual variations.

We make some simplifications and include the developmental effects in two important ways:

1. The maximum firing rates $k_i$ of all populations of cells increases with age.
2. The coupling constants (excitatory and inhibitory) $\alpha_{ij}, \beta_{ij}$ increase with age.

These two assumptions are sufficient to account for developmental changes in the distribution of sleep bouts and wake bouts. The first assumption about increasing firing rates accounts for increasing means of the sleep bouts and the wake bouts. The second assumption accounts for the increasing exponent of the power law in wake bout distribution.

For instance in (8.10), (8.11) and (8.12) we see that initially as mutual excitation parameter increases from 0 to approximately 0.37, all the three populations have only one steady state. Moreover, the firing rate at this steady state does not change even with increasing $\alpha$ for $0 \leq \alpha \leq 0.97$. A bifurcation occurs at $\alpha = 0.37$ and a new stable steady state emerges. The new steady state has a higher firing rate for the wake-active and wake-promoting populations and a lower firing rate for the sleep-active population. Also for $\alpha > 0.37$, the new steady state coordinates $\lambda_1^*$ and $\lambda_3^*$ are increasing functions of $\alpha$. This means that with age the ‘distance’ between the two steady state firing rates is increasing. Heuristically, this is the reason for the increasing exponent in the power law.
Figure 8.7: Wake firing rate ($\lambda_1$) at steady state as a function of $\alpha_{13}$ for $\beta = 0.5$, $\alpha_{31} = 0.5$, $\alpha_{11} = \alpha_{22} = 0$ and $s_1 = 20$, $s_2 = 10$, $s_3 = 10$

Figure 8.8: Sleep firing rate ($\lambda_2$) at steady state as a function of $\alpha_{13}$ for $\beta = 0.5$, $\alpha_{31} = 0.5$, $\alpha_{11} = \alpha_{22} = 0$ and $s_1 = 20$, $s_2 = 10$, $s_3 = 10
Figure 8.9: LC firing rate ($\lambda_3$) at steady state as a function of $\alpha_{13}$ for $\beta = 0.5$, $\alpha_{31} = 0.5$, $\alpha_{11} = \alpha_{22} = 0$ and $s_1 = 20, s_2 = 10, s_3 = 10$

Figure 8.10: Wake firing rate ($\lambda_1$) at steady state as a function of $\alpha := \alpha_{13} = \alpha_{31}$ for $\beta = 0.5$, $\alpha_{11} = \alpha_{22} = 0$ and $s_1 = 20, s_2 = 10, s_3 = 10$
Figure 8.11: Sleep firing rate ($\lambda_2$) at steady state as a function of $\alpha := \alpha_{13} = \alpha_{31}$ for $\beta = 0.5$, $\alpha_{11} = \alpha_{22} = 0$ and $s_1 = 20$, $s_2 = 10$, $s_3 = 10$

Figure 8.12: LC firing rate ($\lambda_3$) at steady state as a function of $\alpha := \alpha_{13} = \alpha_{31}$ for $\beta = 0.5$, $\alpha_{11} = \alpha_{22} = 0$ and $s_1 = 20$, $s_2 = 10$, $s_3 = 10$
9.1 Firing rate plots

In this section, we will assume that the self-excitation parameters $\alpha_{11} = \alpha_{22} = 0$ and the inhibition parameters $\beta_{12} = \beta_{21} = \beta_{32} = 0.5$. The maximum firing rate $s_1$ of the wake-active population is taken to be 20, and the maximum firing rates of sleep-active population and the wake-promoting population $s_2$ and $s_3$ respectively are both taken to be 10. The idea is that under external excitation from LC, the wake-active population’s firing can increase beyond the value that would be obtained simply from self-excitation.

In this section the mutual excitation parameters $\alpha_{13}$ and $\alpha_{31}$ will be assumed to be equal and we will use the symbol $\alpha$ to refer to their common value.

We ran a numerical simulation for 20,000 seconds of model time for the parameters $\alpha_{13} = \alpha_{31} = 0.6$ above and beyond those mentioned earlier. We know from the bifurcation diagram of the deterministic model that for $\alpha = 0.6$ there are two stable fixed points and therefore we expect that the wake firing rate is bimodal.

In figure (9.1) is a short piece of the time course from 300 to 700 seconds. The
wake state is composed of two ‘sub-states’, one where the wake-promoting population (orange) is inactive and one sub-state where it is active.

Figure 9.1: Time course of $\lambda_1$ (in black), $\lambda_2$ (in red) and $\lambda_3$ (in orange); $x$-axis: time in seconds, $y$-axis: firing rate per second

We plot the histograms of the firing rates of the three populations in (9.2) for our numerical simulation. The histograms show that the wake-active and wake-promoting populations have bimodal firing rates. The two modes in the case of the sleep-active population are close so they merge together. Note that for the histogram of the wake-promoting population, we have chosen the maximum $y$-coordinate so that the two modes are clearly visible since the mode located at the low firing rate is higher by several orders of magnitude.
Figure 9.2: Histograms for firing rates of the wake-active population (top two figures), the sleep active population, and the wake-promoting LC for $\alpha_{13} = \alpha_{31} = 0.6$, $\alpha_{11} = \alpha_{22} = 0$, $\beta = 0.5$.

We also ran simulations for the case of non-zero self-excitation, but omit the results since there aren’t any new insights to be gained by studying this case. We only note that as in the two component model we again observe that the sleep and wake firing rates have a stronger self-correlation and a stronger negative correlation with each other.
9.2 Survivor plots

In this section we describe the most important result of our numerical analysis. Recall that we are trying to a model of interacting populations of neurons which gives us desired probability distributions for sleep bouts and wake bouts. Most importantly we are interested in a mechanism for generating a power law distribution for wake bouts. For very young animals the wake bouts are exponentially distributed and as the animal develops, we obtain a power law with an exponent that increases with age.

As explained earlier during the data analysis, a convenient representation used in the biological community is to plot survivor distributions on a semi-log and a log-log scaled graph. We use this representation to demonstrate the numerical results of our model. Recall that when $\alpha = \alpha_{13} = \alpha_{31}$ is made a bifurcation parameter, at approximately $\alpha = 0.5$ a bifurcation occurs and for $\alpha > 0.5$ we have two stable steady states while for $\alpha < 0.5$ there is only one stable steady state. As the animal’s age increases, $\alpha$ increases and eventually we get bistability in the deterministic model.

We have argued that this bistability gives rise to a heavy-tailed distribution of the wake bouts. Heuristically, when the wake bout begins the wake firing rate is close to the lower steady state and as time goes on, it increases and moves closer to the upper steady state. This means that the hazard rate of falling out of wake decreases with time.

We first choose $\alpha = 0.4$ since we know in the deterministic model there is only one
stable fixed point and so we expect wake bouts to be exponentially distributed. This is indeed the case as can be seen in (9.3).

![Survivor plots for sleep (left) and wake (right) on a semi-log scale (top) and a log-log scale (bottom) for $\alpha_{13} = \alpha_{31} = 0.4$, $\alpha_{11} = \alpha_{22} = 0$, $\beta = 0.5$.]

Figure 9.3: Survivor plots for sleep (left) and wake (right) on a semi-log scale (top) and a log-log scale (bottom) for $\alpha_{13} = \alpha_{31} = 0.4$, $\alpha_{11} = \alpha_{22} = 0$, $\beta = 0.5$.

Then we choose $\alpha = 0.6$ which is just after the bifurcation point and since the deterministic model has two stable steady states we expect the wake bout distribution to be a power law which is verified in (9.4). Compare these plots to the ones we get from experimental data in (11.1).
Finally we choose $\alpha = 0.8$ for the survivor plots. This value is close to the next bifurcation where we lose the lower stable steady state. Since the system spends much lesser time close to the lower stable steady state we expect that the wake bout distributions start resembling an exponential distribution again. This is indeed the case in (9.5) especially in the tail of the distribution.
Figure 9.5: Survivor plots for sleep (left) and wake (right) on a semi-log scale (top) and a log-log scale (bottom) for $\alpha_{13} = \alpha_{31} = 0.8$, $\alpha_{11} = \alpha_{22} = 0$, $\beta = 0.5$
CHAPTER 10
MECHANISM FOR POWER LAW

In order to understand the mechanism responsible for generating the power law behavior in the wake bout distributions, we describe a simplified random walk model and relate it to our main model.

10.1 Biased random walk model

In a 2002 paper, C.-C. Lo et. al. [Lo1] gave a random walk model of wake-sleep cycling that is able to generate the correct power law behavior in wake time distributions. By making a connection with our model based on neurological studies, we give an interpretation to the variable that is executing the random walk. On the other hand, the random walk model is significantly simpler since it involves only one dependant variable and so is much easier to analyze mathematically. In fact, we will be able to solve exactly the resulting differential equation.

We first describe the model in Lo et. al. [Lo1]. Let \( X(t) \) be a non-negative time dependant random variable. Let \( h(t) = \mathbb{I}_{(X(t) > 0)} \) be the indicator function denoting the wake or sleep state. By definition, 0 is the sleep state and 1 is the wake state.
Finding the distribution of length of wake bouts is equivalent to finding the first return time of $X(t)$. For $-1 < b < 1$, $X(t)$ evolves according to the following equation

$$X'(t) = -\frac{b}{X} + \xi(t)$$

(10.1)

$$X(0) = X_0 > 0$$

where $\xi(t)$ is a Gaussian white noise with the correlation function

$$\langle \xi_i(t)\xi_j(t') \rangle = 2D\delta(t-t')$$

### 10.2 Connection with the neurophysiological model

The position of the particle in the random walk model corresponds to the firing rate of the wake-active population. Before the wake-promoting Locus Coeruleus becomes active, wake firing rate is centered around a single fixed point. At this early stage in development, during a wake bout both wake and sleep-active populations fire at a constant rate. This means the hazard rate or the rate of transition from wake to sleep remains constant, giving exponentially distributed wake bouts.

After the Locus Coeruleus has become fully active, wake firing rate can be at any of the two deterministic fixed points or anywhere in between. When the wake bout begins, the random walker is close to the lower fixed point. Correspondingly the sleep-active population is at its upper fixed point. As the wake bout progresses, the two firing rates execute a negatively correlated random walk. The inhibition from the
sleep-active population results in the restoring force. As the firing rate of the wake population increases and that of the sleep population decreases, this restoring force decreases. We may model this restoring force as inversely proportional to the wake firing rate.

10.3 Solution of the biased random walk

The general solution of the equation (10.1) can be found in [Br] along with many properties of the solution. Here we give one particular solution and show how it relates to the general solution in [Br].

Again following [Br], we begin by writing down the equivalent Fokker-Planck equation for the probability $P(x, t)$ for a particle to be at position $x$ at time $t$.

$$\frac{\partial P}{\partial t} = D \frac{\partial}{\partial x} \left( \frac{\partial P}{\partial x} + \frac{bP}{x} \right)$$ (10.2)

with the initial condition

$$\lim_{t \to 0} P(x, t) = \delta(x - x_0)$$ (10.3)

where $\delta$ denotes the Dirac delta function. For derivation of the Fokker-Planck equation, see for instance [Ga]. The Fokker-Planck equation is a deterministic Partial Differential Equation for the probability distribution. The trade-off is that it is of a higher order than the original stochastic differential equation.

Since we are interested in the first return time to 0, we demand the Boundary condition
\[ P(0, t) = 0 \quad (10.4) \]

If we ignore the initial condition (10.3) for the moment, we can give a particular solution to the differential equation (10.2).

**Theorem 10.3.1.** A particular solution to (10.2) is given by

\[ S(x, t) = cxt^{-\frac{3+b}{2}} e^{-\frac{x^2}{4Dt}} \quad (10.5) \]

for some constant \( c \).

**Proof:** We will show that (10.5) satisfies (10.2) by explicit calculation, because some of the expressions will be useful later too. The left-hand side of the equation (10.2)

\[ S_t = -\left(\frac{3+b}{2}\right) S + \frac{x^2}{4Dt^2} S \quad (10.6) \]

while the probability current defined by \( j(x, t)(S) = -D \left( \frac{\partial S}{\partial x} + \frac{bS}{x} \right) \) is

\[ j(x, t) = \left( \frac{-D(1+b)}{x} + \frac{x}{2t} \right) S \quad (10.7) \]

So the right hand side of (10.2) is

\[ -\frac{\partial j}{\partial x} = - \left( \frac{D(1+b)}{x^2} + \frac{1}{2t} \right) S - \left( \frac{-D(1+b)}{x} + \frac{x}{2t} \right) \left( \frac{1}{x} - \frac{x}{2Dt} \right) S \quad (10.8) \]

\[ = - \left( \frac{3+b}{2} \right) \frac{S}{t} + \frac{x^2}{4Dt^2} S \quad (10.9) \]

which completes the verification. □
It is easy to see that the solution (10.5) satisfies the boundary condition (10.4). It however does not satisfy the initial condition (10.3). In fact we may look for a solution of the form

\[ P(x, t) = c x t^{-\frac{3+b}{2}} e^{-\frac{x^2}{4Dt}} + O(x^3) \]  

(10.10)

It has been shown by [Br] that

\[ Q(x, t) = x_0^{-\frac{1+b}{2}} x^{\frac{1-b}{2}} \exp \left( -\frac{x^2 + x_0^2}{4Dt} \right) I_\nu \left( \frac{xx_0}{2Dt} \right) \]  

(10.11)

is a solution that satisfies (10.2) and the initial condition (10.3) and the boundary condition (10.4). Here \( I_\nu \) is the modified Bessel function. This solution does indeed have the form \( P(x, t) \).

Since we are only interested in the probability flow \( j(x, t) \) at the origin, we can ignore higher order terms in \( x \) and so the solution \( S(x, t) \) is sufficient for calculating the wake bout distributions.

The first passage time is simply the probability flow through the origin (with a negative sign since the flow is outwards) and so

\[ P_1(t) = -j(0, t) = ct^{-\frac{3+b}{2}} \]  

(10.12)

This is a probability density and so the constant \( c \) can be determined using the normalization condition \( \int_{t_0}^{\infty} P_1(t)dt = 1 \) where \( t_0 \) is a threshold for classification as a wake bout. So that \( c = \frac{1 + b \frac{3+b}{2}}{t_0} \).
The wake bout probability distribution \( P_0(t) \) can be found by integrating \( P_1(t) \) from \( t \) to \( \infty \) for any \( t > t_0 \) and we find that

\[
P_0(t) = \left( \frac{t}{t_0} \right)^{-\frac{1+b}{2}}
\]

(10.13)

**Remark:** We verified both the power law behavior and the exponent in three independent ways.

1. We ran simulations of the system described above. In the simulations the singularity at the origin was regularized by taking a restoring force proportional to \( \frac{1}{x+\epsilon} \) for a small \( \epsilon \).

2. We simulated an equivalent random walk on the positive integers where the probabilities \( p(k) \) of moving left and \( q(k) \) of moving right from the positive integer \( k \) are related by \( p(k) + q(k) = 1 \) and \( p(k) - q(k) = \frac{b}{k+1} \).

3. We numerically calculated the exact probabilities for the random walk described above using a recursion relation. This recursion relation may be thought of as a discrete analog of the Fokker-Planck equation.

In each case we got the same power-law behavior with the predicted exponent.
11.1 Experimental results

How should a scientific experiment be conducted? It is probably not possible to summarize the entire scientific method into a few concise statements. Very broadly speaking however, we can find some commonalities in all experiments. The following steps are perhaps common to most experiments:

1. Design a scientific hypothesis. This can be binary in nature, i.e. a certain statement is true or false, or can range over several possibilities which are ideally mutually exclusive. The aim of the hypothesis and the experiment is to rule out a class of hypotheses and thus strengthen evidence in favor of the remaining hypotheses.

2. Cast the scientific hypothesis in terms of a statistical hypothesis. This involves rewriting the original hypothesis as a null and alternate hypothesis. Usually, but not always, we want to invalidate the null hypothesis.

3. Design a reproducible experiment that will find evidence for or against the
null hypothesis. The experimental design consists of measuring certain pre-
determined variables repeatedly. For convenience of analysis most experiments
are designed so that each measurement comes from an independent and identical
probability distribution.

4. Conduct statistical analysis of the measured observations of a specified variable.

Now I will describe the experimental data that I was working with. The mathematical
model was designed with this experimental setup and data in mind. In mammals,
sleep in fragmented into several sleep and wake bouts. When we think of adult sleep
in humans, we think of sleep as roughly occupying one-third of a 24 hour day usually
in one consolidated 8 hour bout. This view though largely correct, needs refining.
When we investigate closely we find that the consolidated sleep bout in humans is
actually composed of several shorter sleep bouts and several intervening brief wake
bouts. For references see for instance [Blu1], [Lo1] and [Lo2]. This phenomenon
of sleep being composed of micro-awakenings and micro-sleep bouts is even more
pronounced when we look at developing animals, that is animals in their infantile
stage of development.

Sleep bout data has been collected for several different mammalian species, mostly in
adults. The method is to record the time for which each bout, whether sleep or wake
lasts. Then sleep bouts and wake bouts are combined together and we can fit each
to a specified probability distribution. The best fit and the goodness of the best fit is
then recorded. It is universally found across all mammalian species that sleep bouts
follow an exponential distribution while wake bouts follow a heavy-tailed power law.
Blumberg et al. [Blu1] have focused their work on rat infants. The main observation was that when the animal is very young, say around 2 days old, both the sleep and the wake bouts are exponentially distributed (11.1). As the animal develops, that is as it gets older, the sleep bouts remain exponential although the parameter of the exponential distribution changes. The tail of the wake bout on the other hand, gets heavier. In other words, the exponent in the power law increases with age.

For the statistical analysis to be valid, the following three conditions are necessary

1. Wake bouts should be independent from one another.
2. Sleep bouts should be independent from one another.
3. Wake and sleep bouts should be mutually independent.

If we think of a sequence of random variables each measuring the time for which successive bouts last, then the sequence can be thought of as a Markov process trivially since each random variable denoting the bout time does not depend on the previous bout length. This makes it possible to combine all the wake bouts and fit it to a theoretical probability distribution.

For both the wake and sleep bouts, we will plot the complementary cumulative distribution function (c.c.d.f.) also known as survivor distribution, the graph of which is known as the survivor plot \( f(t) := P(T > t) = 1 - P(T \leq t) \). To get some idea about the functional form of the survivor distribution, we plot using a semilog scale (\( y \)-axis is logarithmic, while \( x \)-axis is linear) and a log-log scale (both \( x \)- and \( y \)-axis are logarithmic).
11.2 Distribution fitting

As mentioned earlier, the sleep bouts are well fitted by the exponential distribution for all ages. If $T_s$ is the random variable denoting sleep bout duration then $P(T_s > t) = e^{-\frac{t}{\tau}}$ where $\tau$ is the mean sleep time. The maximum likelihood estimate of $\tau$ is just the mean of the observed sleep bout durations and so it is easy to estimate this parameter from the data.

Before fitting the wake bout distribution, it is necessary to consider a constraint
that is imposed by the experimental setup. We want to find a survivor distribution function which will fit the entire data as best as possible. Due to practical constraints, data is usually not available at all scales, we can only hope for complete data within a certain range. This constraint may not simply be related to difficulties associated with measurement but may be more fundamental. For instance, when classifying a bout as wake, one criterion used in the lab was that the bout must have lasted for at least 1 second. It is not possible to reliably say that the animal is in wake state for a measurement that lasts less than 1 second. A similar remark is true for sleep bouts too. Clearly then the appropriate distribution should be such that the lowest possible bouts are 1 second long. However, even this restriction is not sufficient. The main reason is that bout length measurements for bouts that are just greater than the threshold of 1 second tend to be noisy. The source of the noise is twofold: one is the difficulty of reliably characterizing short bouts as wake or sleep and the second is that the relative error in measurement of bout length is large for smaller bouts. Due to these reasons, the part of the data consisting of smaller bouts may not fit the overall probability distribution as well as the rest of the data.

A possible solution could be to allow for some flexibility in the choice of the threshold, so that we may want to consider only the bouts that exceed a certain minimum bout length. This fitting threshold cannot be less than the measurement threshold of 1 second, but it may be strictly greater in order to reduce the effect of the noisy data. We will look for probability distributions with the property that the form of the distribution is not affected by the choice of the fitting threshold. To state this more
precisely, suppose that $T$ is a one dimensional random variable and $\theta$ is its fitting threshold, which means that measurements of $T$ smaller than $\theta$ are either not possible or are ignored. Choosing a higher $\theta$ allows us to remove the more noise-prone data, the expense however is that we may lose too much data. Note that $\theta$ depends on the measurements of $T$ or to say it another way, it depends on the data. $\theta$ should be chosen such that it is greater than the smallest measurement and such that it achieves a compromise between the two opposing goals of reducing noise while keeping as much data as possible.

Having chosen $\theta$ for a given $T$, we are interested in finding the survivor distribution of $T_\theta := \{T - \theta | T > \theta\}$. Clearly $T_\theta$ is a different random variable than $T$ itself.

What we require is that $T_\theta$ should not be 'qualitatively different' from $T$ for any reasonable value of $\theta$. For instance, suppose that $T$ has an exponential distribution, $P(T > t) = e^{-\lambda t}$. Then from the memory-less property of an exponential distribution, it is easy to see that $T_\theta$ also has an exponential distribution and $P(T_\theta > t) = e^{-\lambda(t-\theta)}$.

So $T_\theta$ has the same mean as $T$, in fact $T_\theta$ is only a translation of $T$.

Clearly we do not expect any other probability distribution to behave as nicely as the exponential distribution under a change of threshold. But there is a class of distribution functions that are used to measure 'exceedances' which have a property that is almost as nice. Without going into a detailed discussion of distributions of exceedances, we mention a particular distribution known as the Generalized Pareto distribution (GP) defined as follows.

**Definition 11.2.1.** Let $t, \sigma, \theta \geq 0$ and let $k \in \mathbb{R}$. Let $T$ be a non-negative valued
random variable. Then $T$ has Generalized Pareto distribution $GP(k, \sigma, \theta)$ if for a fixed $\theta \geq 0$ the complementary cumulative distribution function of $T_\theta := \{T - \theta | T \geq \theta\}$ is

$$P(T_\theta > t) = F_{T_\theta}(t|k, \sigma, \theta) = \left(1 + k \frac{t}{\sigma}\right)^{-\frac{1}{k}}$$

When $k = 0$, the above expression should be viewed in the sense of $\lim_{k \to 0}$. So

$$F_{T_\theta}(t|0, \sigma, \theta) = e^{-\frac{t}{\sigma}}$$

Remarks:

1. When $\theta = 0$, we will drop the third argument and simply write

$$F_{T_\theta}(t|k, \sigma) := F_{T_\theta}(t|k, \sigma, 0)$$

2. $T$ is $GP(k, \sigma, \theta)$ is equivalent to $T_\theta$ is $GP(k, \sigma)$. In other words,

$$F_{T_\theta}(t|k, \sigma, \theta) = F_{T_\theta}(t|k, \sigma)$$

3. The parameters $k$ and $\sigma$ are referred to as shape and scale respectively.

4. For $k = 0$, $\sigma$ is the expected value of $T_\theta$ and correspondingly the mean of the observations of $T_\theta$ gives the maximum likelihood estimate of $\sigma$.

Now we consider what happens under change of threshold.

**Theorem 11.2.2.** Let $T$ be $GP(k, \sigma)$. Then for all $\theta \geq 0$, $T$ is $GP(k, \sigma + k\theta, \theta)$.

**Proof:** $T$ is $GP(k, \sigma)$ is equivalent to saying that $T$ has c.c.d.f. $F_T(t|k, \sigma)$. 

107
\[ P(T - \theta > t|T > \theta) = \frac{P(T - \theta > t \cap T > \theta)}{P(T > \theta)} \]
\[ = \frac{P(T - \theta > t)}{P(T - \theta > t \cap T > \theta)} \]
\[ = \frac{P(T - \theta > t)}{P(T > \theta)} \]
\[ = \frac{P(T > \theta + t)}{P(T > \theta)} \]
\[ = \frac{(1 + k^{t+\theta})^{-\frac{1}{k}}}{(1 + k^{\theta})^{-\frac{1}{k}}} \]
\[ = \left( 1 + \frac{k^t}{1 + k^\theta} \right)^{-\frac{1}{k}} \]
\[ = \left( 1 + \frac{k^t}{\sigma + k^\theta} \right)^{-\frac{1}{k}} \]
\[ = F_T(t|k, \sigma + k\theta, \theta) \]

\[ \square \]

Thus the form of the distribution and the shape parameter are invariant under a change of threshold.

**Note:** Using a simple change of variable argument

\[ F_T(t|k, \sigma, \theta) = P(T - \theta > t|T \geq \theta) = P(T > t + \theta|T \geq \theta) = \left( 1 + \frac{k^t}{\sigma} \right)^{-\frac{1}{k}} \]

we may rewrite the above as

\[ F_T(t - \theta|k, \sigma, \theta) = P(T > t|T \geq \theta) = \left( 1 + k \frac{t - \theta}{\sigma} \right)^{-\frac{1}{k}} \quad (11.1) \]

**Remarks:**

108
1. It has already been observed that the exponential distribution is a special case of GP with $k = 0$.

2. From (11.1) it is easy to see that

$$F_T\left(t - \frac{\sigma}{k}\left| k, \sigma, \frac{\sigma}{k}\right.\right) = \left(\frac{kt}{\sigma}\right)^{-\frac{1}{k}}$$

We define $\theta_0 := \frac{\sigma}{k}$. Then the above tells us that exceedances above $\theta_0$ are distributed according to a usual Pareto distribution, or more simply according to power law. $\theta_0$ thus defines the threshold for onset of power-law like behavior.

These two together show that for special parameter values, exponential and power law are included as special cases of the GP distribution.

3. For $t << \frac{\sigma}{k}$,

$$\left(1 + k\frac{t}{\sigma}\right)^{-\frac{1}{k}} = e^{-\frac{1}{k} \ln (1 + k\frac{t}{\sigma})}$$

$$\sim e^{-\frac{1}{k} (k\frac{t}{\sigma})} = e^{-\frac{t}{\sigma}}$$

For $t >> \frac{\sigma}{k}$,

$$\left(1 + k\frac{t}{\sigma}\right)^{-\frac{1}{k}} \sim \left(k\frac{t}{\sigma}\right)^{-\frac{1}{k}}$$

So we can think of $\frac{\sigma}{k}$ as the transition parameter between exponential and power-law like behavior. In particular, when $k$ is small and non-negative for instance $k = 0$, all the data is exponentially distributed. Conversely, if the smallest data point is greater than $\frac{\sigma}{k}$ then all the data follows a power law.
For the sleep-wake data, the wake bout distribution fits a power law better as age increases because $k$ increases with age.

**Definition 11.2.3.** The hazard function or hazard rate is defined to be the ratio of the probability density function to the survivor function. In particular, if the survivor function $S(t)$ is differentiable then

$$h(t) = \frac{-S'(t)}{S(t)} = -(\ln(S(t)))'$$

**Theorem 11.2.4.** The hazard function of $GP(k, \sigma)$ is

$$h(t) = \frac{1}{\sigma + kt}$$

*Proof:* Trivial calculation. □

*Remarks:*

1. An interpretation of the hazard rate is the instantaneous rate of transition at time $t$ conditioned on the event that no transition has happened up to time $t$.

2. In particular for the exponential $k = 0$ case, the hazard rate is constant in time and in fact the hazard is just the reciprocal of the expected value of the random variable $h(t) = \frac{1}{\sigma}$.

3. For $k > 0$, $h(t)$ is a decreasing function of $t$. $\frac{1}{\sigma}$ gives the initial hazard when the bout begins. For a larger value of $k$, the hazard decreases faster with time.
11.3 Parameter estimation

In the previous section, we gave theoretical arguments in favor of fitting the wake bout distributions using a Generalized Pareto and the sleep bout distributions using an exponential. In this section we show that these distributions do indeed constitute a good model. To achieve this, we need to do the following:

1. Show that GP and exponential are the best fit models among the candidates.

2. Using goodness-of-fit tests, show that the best fit model is also a good fit.

3. Estimate parameters for the model.

At first glance even though the above steps may seem to be organized in a logical order, it is not possible to execute them in that order. For instance, in order to compare two forms of distributions in step 1, say GP with another distribution $X$ we need to estimate the parameters first (step 3). So we need to find a process or algorithm that executes all the steps, maybe in a different order.

I will now describe the numerical scheme that I used to estimate the parameters, to find the best fit model and to evaluate the goodness of fits. The general framework used is that of Maximum Likelihood. The likelihood function $L(\text{parameters}|\text{data}) \propto P(\text{data}|\text{parameters})$ will play an important role. The likelihood function allows us to find the values of the distribution parameters which are most likely to result in the data that we have observed.

The method of parameter estimation that is used here is similar to the one proposed by Clauset et al [CSN]. The procedure outlined by Clauset et al is for the two
parameter power law. Since GP has three parameters we use a modified version of the method as explained in the following steps.

1. Choose a family of distributions that is likely to fit the data. The family is parametrized by a set of parameters, for instance \( \{k, \sigma, \theta\} \) for the GP distribution.

2. **Parameter estimation**: Find the parameter values within this family that give the best fit to the data. We do this by calculating the likelihood function and then finding the maximum of the likelihood function.

3. Decide the class of candidate distributions. For instance, for theoretical reasons, we may decide that the candidate distributions are GP, exponential, Pareto, and Weibull.

4. In each of the above families, we can find the parameters that maximize the likelihood function. Now we can compare across families to find which family maximizes the likelihood function.

5. Having found the appropriate family of distribution and the best estimates of parameter values, we now determine if the best fit is actually a good fit. We use Kolmorov-Smirnov (KS) test or a slightly modified version of Kolmogorov-Smirnov test to do this. The KS test is also used as part of parameter estimation in step 2 as will be explained below.
11.3.1 Parameter estimation and likelihood function for GPD

Given a threshold \( \theta \), several numerical schemes have been developed which allow for estimation of \( k \) and \( \sigma \). For references, see for instance, [Gr], [Ca], [Em] and [Ko]. None of these numerical schemes estimate \( \theta \). There are usually two reasons that are provided for fixing a threshold in advance and not estimating it from data, one reason is theoretical and the other is practical. The theoretical reason is that the threshold is something that is imposed by the experimental (or measurement) process rather than a feature of the data itself. For instance, we may be interested only in flooding or rainfall beyond a certain threshold value. The practical reason is that there is no Maximum Likelihood Estimate for the threshold parameter \( \theta \).

We have already noted earlier that we are interested in finding the fitting threshold \( \theta \) which is greater than or equal to the experimental threshold \( \theta' \). This means that we have to estimate \( \theta \) from the data. Since there is no MLE for \( \theta \), we will use the Kolmogorov-Smirnov test for finding the best estimate for \( \theta \). We will usually impose some maximum cut-off (say 3 seconds) on the value of \( \theta \), since if we allow the threshold to take very large values, even though the fit becomes better it is at the expense of losing significant amount of data.

**Algorithm for estimating the threshold value \( \theta_c \)**

1. Fix \( \theta \in [\theta_{\text{min}}, \theta_{\text{max}}] \).

2. Based on this value of \( \theta \) and from the data, estimate the other parameters \( k \) and \( \sigma \) using maximum likelihood methods.
3. Calculate the theoretical probability distribution function $F(t|k, \sigma, \theta)$.

4. From the data $T_i, 1 \leq i \leq n$, calculate the empirical distribution function

$$F_n(t) = \frac{1}{n} \sum_{i=1}^{n} I_{T_i \leq t}$$

where $I$ is the indicator function.

5. Calculate the Kolmogorov-Smirnov (KS) statistic defined as

$$D_n(\theta) = \sup_t |F_n(t) - F(t)|$$

6. Repeat the previous steps by varying the threshold $\theta$ and choose $\theta_c$ to be such that

$$\theta_c = \{\theta | \min_{\theta \in [\theta_{\min}, \theta_{\max}]} D_n(\theta) \}$$

i.e. $\theta_c$ is the value of $\theta$ such that the KS statistic is minimized.

**Note:** The data near the tails of the distribution is weighed more than the middle in calculating the KS statistic. To remedy this, sometimes the modified KS statistic is used. The modified KS (MKS) statistic is defined as

$$K_n(\theta) = \sup_t \frac{|F_n(t) - F(t)|}{\sqrt{F(t)(1 - F(t))}}$$

We used both KS and MKS statistic. Generally the MKS statistic was preferred and gave better results. One consistent advantage of using MKS statistic was that the value of $\theta_c$ thus calculated was usually lower than that calculated using the KS statistic. This is desirable because it results in less loss of data.
11.4 Results

1. We consistently find that sleep data at all ages, for all individual animals and for the pooled data is fit better by the exponential distribution. Even though the fit for sleep data improves when we use GP (recall that exponential is a special case of GP), the improvement in fit is not enough to justify including an additional parameter.

More precisely, Exponential is considered to be a nested hypothesis within the framework of GP. We used Vuong test, which is a modification of the likelihood ratio test for nested hypothesis to test if the extra parameter introduced by GP was necessary. For theoretical details about the test please refer to [Vu].

2. GPD provided a better fit for wake data at all ages when compared with Exponential, Pareto, Weibull (or stretched exponential), and Generalized Extreme Value distribution.

3. The shape parameter $k$ increased monotonically with age, $k$ was close to 0 for animals aged 2 days and $k$ was close to 1 for animals aged 21 days.
Appendix A

POINCARE-BENDIXSON THEOREM

For $x \in \mathbb{R}^n$ consider the system $x' = f(x)$. Let $\phi(t, x)$ denote the flow generated by $x' = f(x)$.

**Definition A.0.1.** A point $x_0 \in \mathbb{R}^n$ is called an $\omega$ limit point of $x \in \mathbb{R}^n$ if there exists a sequence $\{t_i\}, t_i \to \infty$, such that

$$\phi(t_i, x) \to x_0$$

$\alpha$ limit points are defined similarly by taking the sequence $\{t_i\}$ as $t_i \to -\infty$.

**Definition A.0.2.** The set of all the $\omega$ limit points ($\alpha$ limit points) of a flow is called the $\omega$ limit set ($\alpha$ limit set), denoted $\omega(x)$ ($\alpha(x)$).

**Theorem A.0.3.** (Poincare-Bendixson) Let $\mathcal{M}$ be an invariant vector region of the vector field containing a finite number of fixed points. Let $p \in \mathcal{M}$, and consider the $\omega$ limit set of $p$. Then one of the following possibilities holds.

1. $\omega(p)$ is a fixed point.
2. $\omega(p)$ is a closed orbit.
3. $\omega(p)$ consists of a finite number of fixed points $p_1, \ldots, p_n$ and orbits $\gamma$ with $\alpha(\gamma) = p_i$ and $\omega(\gamma) = p_j$.
In other words, in a plane if a trajectory is confined to a compact set then it can only do one of the following: converge to a fixed point, be on a limit cycle or converge to a limit cycle.

The main application of this theorem will be to show that all trajectories converge to fixed point solutions and if there is a unique fixed point then it is globally asymptotically stable. The next appendix helps us to rule out periodic orbit solutions.
Appendix B

MONOTONE DYNAMICAL SYSTEMS

A complete treatment of the subject of Monotone Dynamical Systems is beyond the scope of this dissertation. The reader interested in more details than are presented here should consult the references [Hir], [Son1] and [Son2], for instance. Rather than give definitions of monotone systems in full generality, we focus on the applications of the main results of the theory. Most of this appendix is based on the review paper [Son1].

The appeal of monotone systems in biological modeling comes largely from two important properties.

1. Monotone systems are robust, in the sense that they have predictable and constant response to inputs for a large range of parameters.

2. Complex biological systems may sometimes be decomposed into monotone subsystems where each subsystem has predictable behavior.

To quote Sontag [Son1] “Monotone subsystems have appealing properties as components of larger networks, since they exhibit robust dynamical stability and predictability of responses to perturbations. This suggests that natural biological systems may
have evolved to be, if not monotone, at least close to monotone in the sense of being decomposable into a “small” number of monotone systems.”

One of the main applications of monotone system theory is to rule out periodic orbits or chaotic behavior. In some sense, monotone systems behave a lot like one-dimensional systems (in fact, all one-dimensional systems are monotone) and therefore all bounded trajectories converge to fixed point solutions.

We construct a graph $G$ of a dynamical system $x'_i = f_i(x_1, \ldots, x_n), i = 1, \ldots, n$ according to the following rules.

1. If $\frac{\partial f_i}{\partial x_j}$ is non-negative then we draw an excitatory arrow $\rightarrow$ from $j$ to $i$.

2. If $\frac{\partial f_i}{\partial x_j}$ is non-positive then we draw an inhibitory arrow $\sqsubset$ from $j$ to $i$.

3. If the sign can be positive or negative then we draw an arrow from $j$ to $i$ with a sign $\pm$.

4. If the derivative is identically zero, then we don’t draw an arrow.

Note that such a graph corresponds exactly to the graphical representation for the wake-sleep system discussed in (3.6)

**Theorem B.0.4.** *(Hirsch)* Suppose that a dynamical system satisfies the following properties.

1. The graph $G$ is strongly connected, which means that the directed graph $G$ has a path between any two vertices.

2. Every non-identically zero Jacobian entry is everywhere non-zero.
Then for every initial condition, except perhaps for a measure-zero set of initial conditions, all bounded solutions of the dynamical system converge to a set of steady states.

Both the conditions are quite easily checked for the wake-sleep system. From figure (3.3), we can easily see that the graph for the wake-sleep system is strongly connected. The second criterion is also easily checked since \( \frac{\partial h_i}{\partial \lambda_j} = g_{ij} (\lambda_i) \) which is either strictly positive if \( g_{ij} \) is the excitation coefficient or strictly negative if \( g_{ij} \) is the inhibition coefficient for all values of \( \Lambda \in (0, s_1) \times (0, s_2) \times (0, s_3) \).

A more graphical approach to checking the second criterion is to multiply signs. Attach a positive sign for every excitatory connection and a negative sign for every inhibitory connection. Assume that the graph is undirected. Multiply all signs along any closed path. If the product is strictly positive then the system is strongly monotone. This graphical approach tells us about which components of a system might be monotone at a glance. We can also imagine embedding the wake-sleep system or any monotone system within a larger system which has a more complex range of behaviors.

A final remark about monotone systems is that any subsystem of a monotone system is also monotone. In particular the reduced two component system comprising the wake-active and sleep-active populations is also monotone.
BIBLIOGRAPHY


[Si] Jerome M. Siegel, Clues to the function of mammalian sleep, Nature Insight, 437, (October 2005), pp. 1264-1271
