Modelling Biological Clocks: A Study of the Repressilator

Somdeb Ghose

The Institute of Mathematical Sciences
CIT Campus, Taramani, Chennai 600 113, India

Thesis submitted in partial fulfilment of the requirements for the award of the degree of

Master of Science

HOMI BHABHA NATIONAL INSTITUTE

May 2008

Abstract

Biological clocks are found everywhere in the natural world. From circadian (day-night) rhythms to pacemakers of the heart, biological clocks are an essential part of the smooth functioning of living organisms. They are composed of microscopic clocks which operate at the cellular level. The macroscopic clock functions by synchronisation of these smaller units through intercell coupling. These clocks keep time with amazing regularity even in the face of random external fluctuating effects or internal noise which are ubiquitous in the natural world. In order to attempt to understand the robustness of such clocks against stochastic effects, a model genetic circuit called the repressilator was constructed as a simple clock. The motivation is to find generic features which underlie the operation of complex biological clocks in a simple model. In this thesis, we review previous work done on single repressilators as well as on repressilators coupled by the quorum sensing mechanism. We perform numerical simulations to study the properties of a mathematical model of the repressilator. We suggest a new mechanism by which repressilator circuits can be coupled. Numerical simulations of two repressilators, coupled by our mechanism, show attractive phase synchronisation when their natural frequencies are identical. When the natural frequencies are distinct, our results rule out entrainment and seem to indicate the absence of phase-locking.

Acknowledgment

I would like to thank Dr Ronojoy Adhikari for introducing me to this problem and guiding me through every difficulty faced while attempting to understand and study the wonderful workings of the repressilator, as well as for reviewing this thesis and offering critical and helpful suggestions. Working with him has instilled in me a deep liking for—and interest in—this field. He has been—and continues to be—inspirational, and I am grateful to him for that.

Sandeep K Goyal is a colleague and a senior research fellow in Physics at IMSc. Discussions with him have been most useful regarding this problem, as have his suggestions and advice which have motivated me and resolved the occasional deadlock. I thank him.

I would like to thank Madhuvanthi Raj Iyer of SHASTRA college, Tamil Nadu, India for detailed discussions on this topic as well as for help with a few of the biological aspects of this problem.

Rajeev Singh is a fellow researcher in Physics at IMSc. Discussions with him have resolved certain issues regarding the concept of synchronisation. He has also been helpful in suggesting literature which have been most useful in studying this problem. I thank him too.

Finally I would like to thank Neeldhara Misra, Rajeev Singh and Sandeep K Goyal for help with LaTeX, Neeldhara Misra for help with the GIMP image editor and Anoop Varghese and Neeldhara Misra and Ronojoy Adhikari for enduring mock-talks concerned with this material.

Contents

List of Figures

Chapter 1

Introduction

The Dutch physicist Christiaan Huygens is most well known for his discovery of the pendulum clock. He is less well known as the discoverer of a rather interesting and curious phenomenon pervading in pendulum clocks. The year is 1665, the month is a cold February, and Huygens is confined to his bed with some small ailment. With nothing better to do, he falls into an idle study of the two pendulum clocks hanging side by side on the wall—clocks he had recently built. Curiously, he observes that the clocks are running in exact synchrony. This synchronisation lasts for as long as he watches the clocks. Intrigued, he gets up and deliberately desynchronises them. To his astonishment, they fall in step once again after some time and stay in mutual synchrony. Huygens then tries another approach. He shifts one clock to another part of the room—the wall opposite to the first clock. This time the two clocks fail to march together—Huygens observes one losing about five seconds per day relative to the other. Thus he concludes that the two clocks synchronise through tiny coupling forces transmitted through air or the wall to which they were attached[12].

Huygens' accidental observation has created what is now known as the theory of coupled oscillators[1]. Such coupled oscillators are ubiquitous in the natural world. In fact, every human being has a proliferation of such coupled oscillators. Pacemaker cells in the heart, insulin-secreting cells in the pancreas, neural networks in the brain and spinal cord are all examples. These collections of oscillators perform such diverse and fundamental physiological functions such as pumping blood throughout the body or regulating the level of hormones in

[1] An oscillator is any system that executes periodic behaviour.

the blood. They enable the performance of basic jobs such as breathing, running or chewing[12]. It is also very much possible that not all biological coupled oscillators need to be confined within the same organism. A superb example of this is in congregations of synchronously flashing fireflies. Another example is a

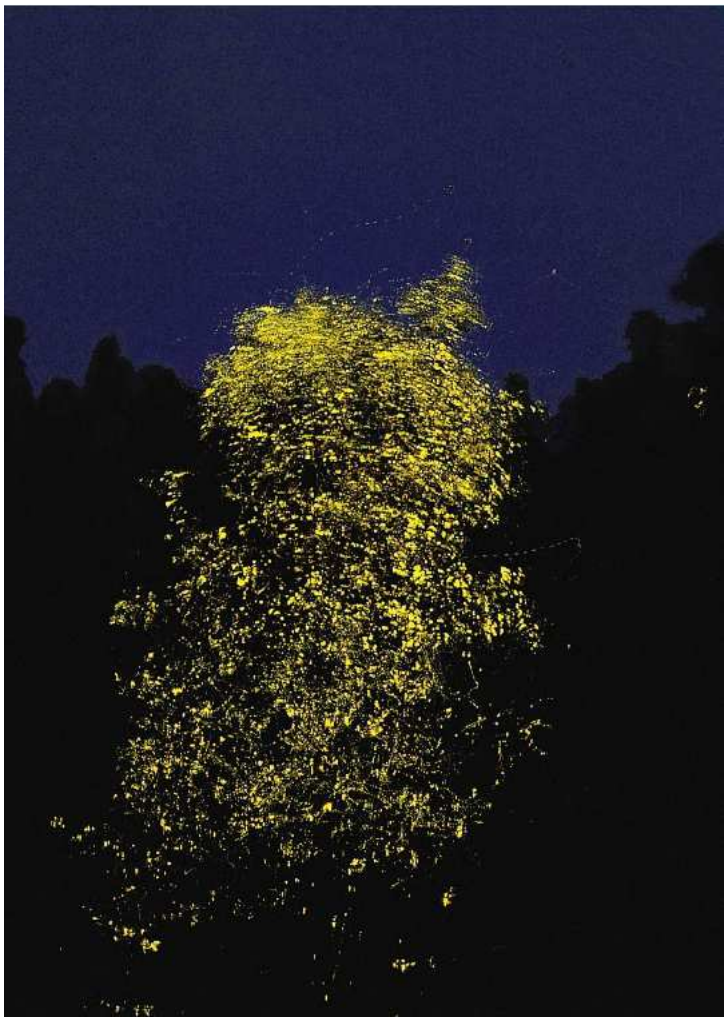

Figure 1.1: Thousands of fireflies flash in unison in this time-exposure of a nocturnal mating display. Each insect has its own rhythm, but they synchronise with the others through its nearest neighbours' rhythms. Picture from the paper by Strogatz and Stewart[12]

group of crickets that chirp in absolute (and sometimes maddening) synchrony.

In order to understand how coupled oscillators work together, it is important to understand how a single oscillator operates. A pendulum is a good example of a single harmonic oscillator. A harmonic oscillator is a linear oscillator.

Its amplitude is fixed by initial conditions while its frequency can be varied by changing the values of the parameters (the length of the pendulum, for example). Amplitude and frequency can be tuned independent of each other for such oscillators. However, a biological oscillator is a non-linear oscillator and its amplitude and frequency cannot be varied indepent of each other. It is due to non-linearity that it is necessary to thoroughly understand the mechanism by which a single oscillator operates.

Coupled oscillators can synchronise in a variety of possible ways. In the case of two coupled oscillators, phase-locking[2] can be brought about in two ways—synchrony and anti-synchrony[12]. In the former case, the phase difference between the oscillations of the two oscillators are completely in phase, *i.e.* their relative phase is zero. In the latter case, the relative phase difference is π. These two phenomena are also known as attractive and repulsive entrainment[3][5]. It is to be remembered however that if two oscillations are phase locked it does not mean they are synchronised. Only if the phase difference between the two oscillators is 0 or π can they be called synchronised. If the phase difference should fall between these two values, the oscillations are simply phase-locked, and not synchronised in the technical sense[12][5]. However, this is a matter of convention. Other authors[10] use the words 'phase-locking' and 'synchronisation' equivalently. Perhaps it would be better to say that there is 'phase-synchronisation' if there is 'phase-locking'[8]. In this thesis, we have adopted the Kuramoto and Strogatz viewpoint and have meant 'complete synchronisation', *i.e.* phase-locking by 0 or π when we speak of 'synchronisation'.

A nice example of attractive and repulsive coupling in the natural world is kangaroos and humans running[12]. The hind legs of kangaroos fall in step as it runs, showing attractive coupling, whereas a human runs on alternating legs displaying repulsive coupling. Three or more coupled oscillators show even more diverse behaviour. Three oscillators can couple in four ways: all in phase; all out of phase by $\pi/3$; two in phase and the third out of phase by any value and finally two out of phase by π and the third going twice as fast. The final pattern is exhibited by a man walking slowly with the help of a walking stick. His legs have

[2] A definite phase difference existing between oscillators with the same patter of oscillation

[3] Entrainment is a term used by Yoshiki Kuramoto that means the same thing as synchronisation in the sense we use it here, which is phase synchronisation with a phase-difference of 0 or π.

a π phase difference, while his walking stick has a frequency double that of his legs (see figure 1.2). Another important class of oscillatory phenomena shared by many living organisms in the biological world are circadian rhythms[1][5]. Circadian rhythms are approximately 24-hour cycles in the biochemical, physiological or behavioural processes of living beings. Under normal environmental conditions, such oscillations are synchronised by periodic external forces associated with the sunrise-sunset cycle. However, it has been observed that such oscillations continue taking place (albeit with a slight deviation form their 24-hour cycle) even if the periodic driving forces are removed. This leads to the conclusion that there are limit-cycle oscillators in-built into living organism. In the microscopic domain, these consist of a collection of mutually coupled and synchronised (or at least phase-locked) cellular oscillators.

A very interesting aspect of synchronisation among coupled oscillators is that the onset of synchrony can be treated as a phase-transition[4][5]. In natural systems, a collection of coupled oscillators are very seldom exactly identical—there remain some (small) difference between them, mostly in the form of a range of natural frequencies for the collection. Additionally, environmental effects are ever-present, and these effects can introduce different responses from otherwise identical oscillators. Thus such oscillators affected by random external effects form a statisticl ensemble. The external stochastic perturbations can be detrimental to the attainment of mutual entrainment[5]. Contrary to this, coupled oscillators have the tendency to achieve mutual entrainment. These two opposing effects play against each other in driving the phase transition of the system from a non-entrained to an entrained state[5].

Synchronisation is controlled by two independent aspects—the properties of the single oscillator and the number of oscillators that are coupled. Stochasticity, both internal and external, affect the time-keeping ability—or the stability of oscillation—of a single oscillator. It is primarily the number of oscillators which determine the time-keeping ability—the stability of synchronisation—of a collection of oscillators. We use a master equation approach in order to account for internal noise, external random influences and the effect of varying number of oscillators. In the limit of small fluctuations we recover the deterministic

[4] A rapid change of a system from one state to another for a very small change in control parameter(s)

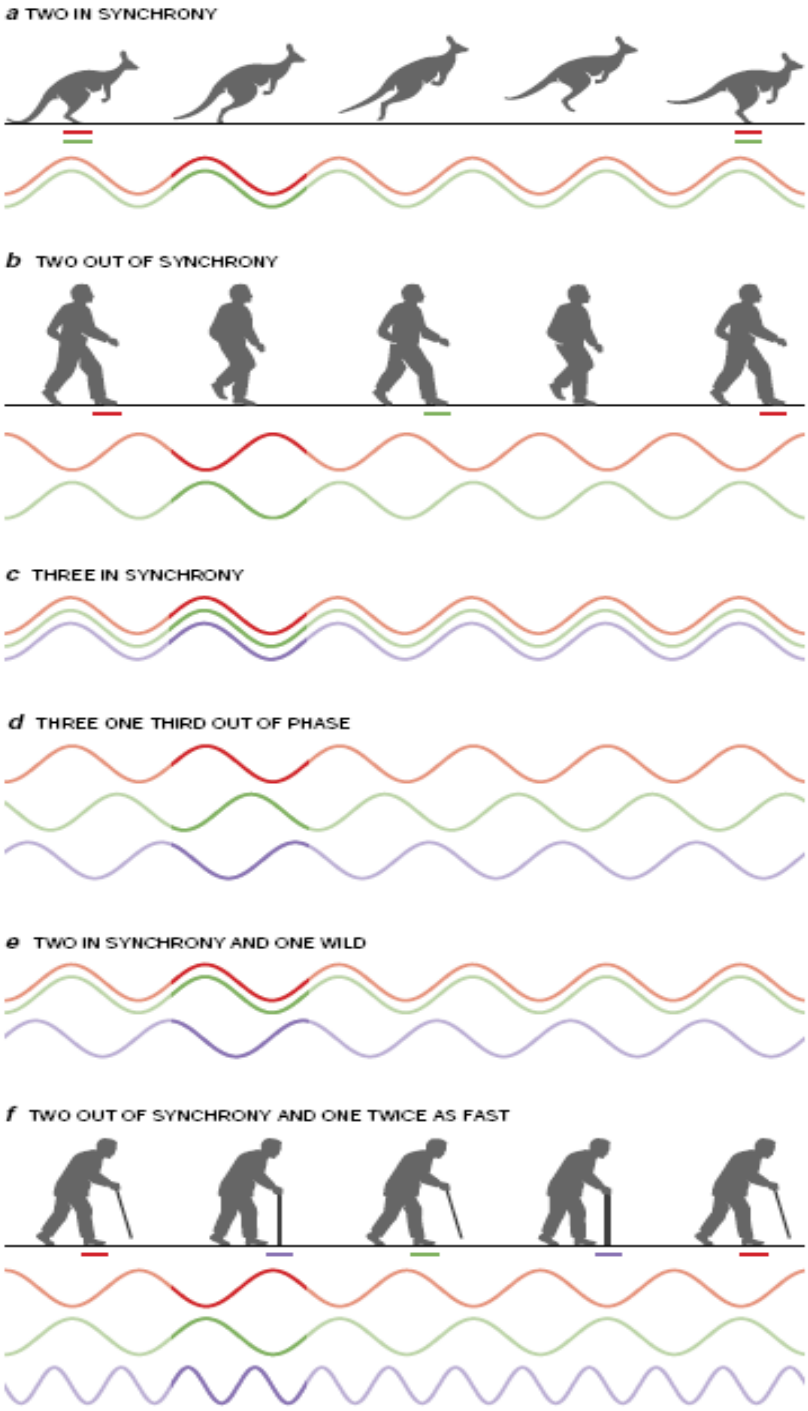

Figure 1.2: Different examples of phase-locking in the natural world—man and kangaroo. Figure from the paper by Strogatz and Stewart[12]

description.

The outline of the thesis is as follows. We have studied the repressilator which a simple model clock constructed using a negative feedback genetic circuit. The motivation is to find generic features which underlie the operation of complex biological clocks in a simple model. We have reviewed previous work done on single repressilators as well as on repressilators coupled by the quorum sensing mechanism. Chapter 2 contains a description of the repressilator as well as a review of experimental results regarding the operation of the repressilator. Chapter 3 reviews the mathematical models (both deterministic and stochastic) used to model the repressilator as well as computer simulation results of those models. Chapter 4 contains our analysis of the repressilator. We have performed numerical simulations to study the properties of existing deterministic and stochastic mathematical models of the repressilator. We have also suggested a new mechanism by which repressilator circuits can be coupled and have performed numerical simulations of two repressilators coupled by our mechanism in order to study entrainment between the two oscillators.

Chapter 2

The Repressilator

2.1 Introduction

Sailing ships in the 16th and 17th century faced a singularly difficult navigational problem. Navigators in the Age of Sail were experts in finding out the current latitude of their ship, but it was the accurate determination of the current longitude that foxed them. One solution was to have two clocks on board—one kept at Greenwich time and the other kept at ship time according to the positioning of the sun at zenith, which was marked as noon. The difference in time between the two clocks would give the current longitude. However, differences in temperature and rough sea conditions played havoc with the accurate time-keeping abilities of the two clocks. Errors in determining longitude, however small time-wise, would translate into big differences in actual distance. For instance, one degree longitude approximately equals 111 km at the equator. One degree is also equivalent to four minutes of time ($24 \times 60/360$). So an error of a mere four minutes in timekeeping would translate into an error of 111 km or 69 miles or 60 nautical miles. This might seem small in the vast oceans, but this error tended to accumulate. Consequentially, people were lost at sea. The need of the hour was to build an accurate and robust clock which would keep perfect time in face of even a thundering hurricane. The British government offered an incredible £20,000 (about £6 million at present) for the building of this clock. John Harrison was the man who did it[11].

The point of this is that John Harrison was able to make a clock that kept accurate time even in the face of severe weather conditions like rolling seas and

temperature differences. The clock is an oscillator, and the weather conditions are stochastic perturbations that tend to force the clock to run incorrectly. A similar problem is faced by biological oscillators. Stochasticity is inherent in the natural world, and noise is a constant hindrance to the suitable working of any biological oscillator. However, in spite of this, the familiar biological oscillators in the brain or the pacemaker of the heart are incredibly robust and keep accurate time. If they had been unable to do so, life as we know it might not have survived. So the question is, what is it that lends robustness to these natural oscillators? To study this question, physicists have made a (sort of a) 'spherical cow' model—a simple clock, constructed out of three genes connected to each other in a negative feedback circuit. This is the Repressilator. It was Michael Elowitz and Stanlislas Leibler[3] who first proposed and constructed such a model in the bacterium *Escherichia coli*.

2.2 Description of a repressilator

2.2.1 Generic Schematic Description

A gene is a segment of DNA[1] that codes for a particular protein or RNA[2]. It is the smallest unit of an organism that is still able to contain and transfer genetic information. The genes, located on the chromosomes, contain the information for the production of proteins. Gene expression is a process through which a gene produces protein, and consists of two sub-processes: *Trascription* and *Translation*. *Transcription* is the production of messenger RNA from the genes, while *translation* is the conversion of mRNA into protein.

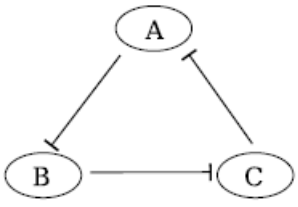

Figure 2.1: Schematic description of a repressilator (figure from [3])

[1] DNA: deoxyribonucleic acid
[2] RNA: ribonucleic acid

Figure 2.1 describes the repressilator schematically. A repressilator has three genes a, b and c along with the corresponding proteins they express: A, B and C. These proteins act as repressors, *i.e.* they repress or inhibit the rate of expression of the gene next in line. For example, protein A represses gene b, slowing down the production of protein B. This inhibitory effect increases as the concentration of the protein increases. Similar to A repressing b, B represses c and finally C represses a to complete the cycle. Hence it is a negative feedback loop. For example, if we were to start out with a bit of A and no B or C, and if these genes were *not* connected, then the concentrations of each of A, B and C would keep on increasing since the rate of expression is constant. The only check to this would then be the natural degradation of the proteins (maybe into their constituent amino acids), which is also always present at a constant rate. However, once they are connected, a high concentration of A would decrease the expression rate of b, which would decrease the concentration of B. This would then let c express at a faster rate, which would increase the concentration of C. This would ultimately result in a checking effect on the production of A, which would have to decrease. Then b would be freer and B would increase, decreasing C and increasing A again. Hence, the cycle would start once again. Oscillations can be expected, and are found, in the concentrations of the proteins in the steady state (*i.e.* after a transient state).

2.2.2 Elowitz and Leibler's construction

The repressilator created by Elowitz and Leibler is described thus (see figure 2.2). TetR, λcI and LacI are the three proteins that are expressed by their parent genes *tetR*, *cI* and *lacI*. The gene *lacI* is obtained from *E. coli*, *tetR* from the tetracycline-resistance transposon[3] Tn*10* and *cI* from λ phage. The protein LacI inhibits the expression of the gene *tetR*. The protein TetR in turn inhibits the expression of the gene *cI* while protein λcI inhibits the expression of *lacI*, thus completing the cycle. So it is a negative feedback circuit, which could lead to oscillations, as we shall see. A green fluorescent protein (GFP) is used to detect the concentration levels of the repressilator components. In the second figure it can be seen that TetR inhibits GFP, so that decrease in intensity of

[3] A transposon is a discrete piece of DNA that can insert itself into other DNA sequences within the cell

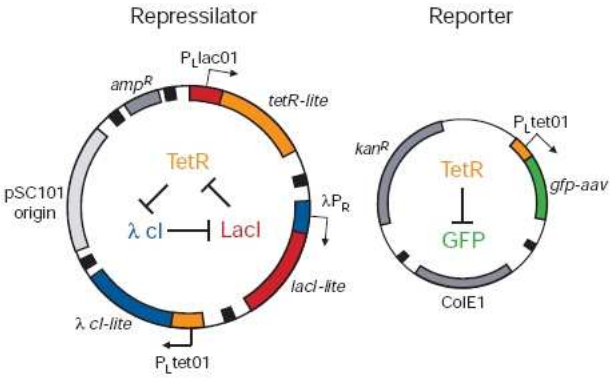

Figure 2.2: Diagram of the repressilator showing the genes and the proteins that make up the repressilator (figure from [6])

the GFP as seen by an optical microscope indicates high concentration of TetR. The experiment performed by Elowitz and Leibler showed oscillations in the concentrations of the proteins with time periods greater than cell-division time, implying that the state of the oscillator is transmitted through generations.

2.3 Experimental Results: Elowitz and Leibler

Elowitz and Leibler built their repressilator in the bacteria *E. coli*, and used GFP to observe the presence (or absence) of oscillations. Since single cells had no apparent means to achieve or maintain synchrony, individual cells were isolated under the microscope and the their fluorescent intensity was studied as these cells grew into small two-dimensional micro-colonies consisting of hundreds of progeny cells. The graph shows temporal oscillations in GFP fluorescence intensity with a time-period of roughly 150 minutes, which happened to be almost three times as large as cell-division time-scales. Snapshots of a microcolony of bacteria are given in figure 2.3, while the fluorescence intensity of the marked cell against time is given in figure 2.4.

Looking at the graph (figure 2.4), one can see that the time-period of oscillations is around 150 mins (peak-to-peak). The bars at the bottom of the graph show that septation[4] occurs about three times per oscillation cycle on average, that is, with a time period of about 50 mins. This conforms to standard

[4] Septation means cell-division

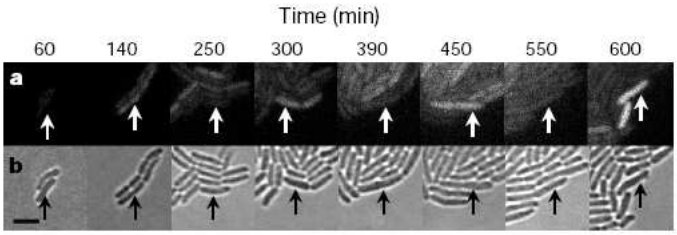

Figure 2.3: Snapshots of a microcolony of *E. coli* taken in (a) fluorescence and (b) bright-field. The arrows indicate the cell under observation (figure from [3]).

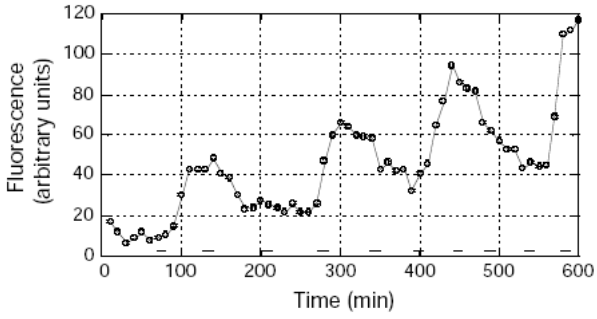

Figure 2.4: Oscillations in GFP fluorescence intensity (bars at the bottom indicate septation events) (figure from [3])

cell-division time-periods. Therefore it is seen that the state of the network is transmitted to its offsprings despite there being noise in the form of stochastic fluctuations in the dynamics of the cellular clock. However, significant differences in period and amplitude of the oscillator were observed in the oscillator output. These differences were seen among cells in different lineages (descended from different ancestors) or among the cells in one lineage, which are called siblings (see figure 2.5).

Elowitz and Leibler conclude that it is possible to design and construct a new artificial genetic oscillator with new functional properties from generic components. It might then be possible to attempt to understand the design principles of more complex oscillators such as the circadian oscillator using the simple repressilator model. However, as opposed to the robust behaviour of the circadian clock, the behaviour of the repressilator (according to the these experimental results) seemed to be noisy and variable.

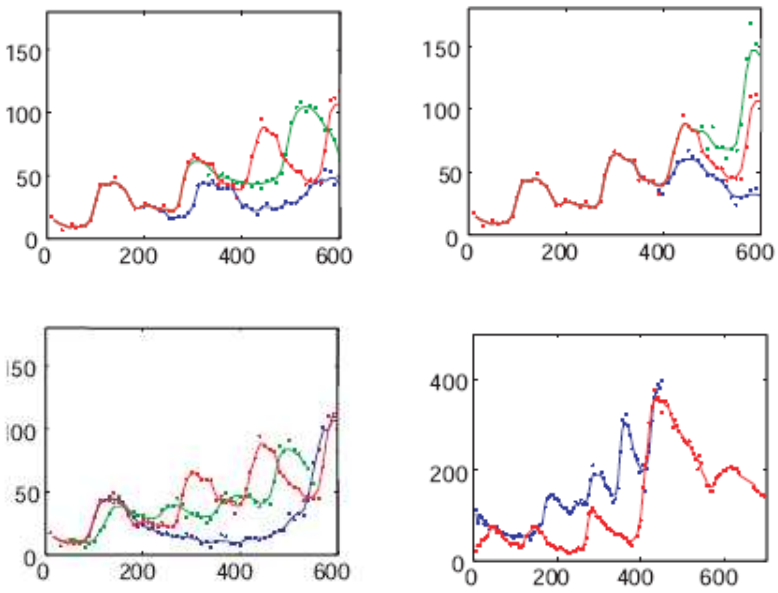

Figure 2.5: Top left figure shows phase delays after septation among sibling cells (blue and green) relative to the reference cell(red). Top right figure shows that phase is maintained but amplitude varies greatly after septation. Bottom left figure is for reduced period (green) and long delay(blue). Bottom right figure shows large variations in period and amplitude. The top two figures and the bottom left figure show the variation of oscillation among cells of the same lineage. The bottom right one shows the variations between cells of different lineages. Figures from [3].

2.4 Quorum Sensing

Elowitz and Leibler's results show that there are significant variation in frequency and amplitude among different repressilators. The problem with using a small number of biological oscillators is that stochastic fluctuations (noise) tend to dominate the oscillations preventing mutual entrainment or effective synchronisation. Intercell coupling is a way to address these issues. But how would one couple different cells (repressilators)? A possible answer is through quorum sensing.

2.4.1 What is Quorum Sensing?

Bacteria can communicate with each other by releasing and detecting 'signalling molecules' within their environment. The term 'quorum sensing' is used to describe the phenomenon whereby the accumulation of signalling molecules enable a single cell to sense the number of bacteria (cell density). However, in the natural environment, different classes of bacteria may use different signalling molecules, so that communication is possible only within the same class of bacteria.

Quorum sensing allows bacteria to communicate with each other and co-ordinate their behaviour. Conditions in the natural world change often and quickly, and quorum sensing allows bacteria to respond speedily and adapt themselves to the change in order to survive. For example, discovery of a source of nutrients needs to be communicated quickly to the entire bacterial population for first claim. Also, conflict with other organisms might be necessary for securing claim over the source of nutrients. Co-ordination of behaviour (timing aggression and defence) is done using quorum sensing. Again chemicals like antibiotics are toxic to the bacteria. Detection and avoidance of regions with high toxicity concentration (for the bacteria) is essential for survival. Additionally, it is very important for pathogenic[5] bacteria during infection of a host (e.g. humans, other animals or plants) to co-ordinate their virulence in order to escape the immune response of the host and establish a successful infection.

[5] A pathogen (from Greek pathos, suffering/emotion, and gene, to give birth to), infectious agent, or more commonly germ, is a biological agent that causes disease or illness to its host.

2.4.2 Using Quorum Sensing in Repressilators

Jordi Garcia-Ojalvo, Michael Elowitz and Steven Strogatz[4] have used the idea of quorum sensing to couple different non-identical and noisy repressilators. They have shown through computational modelling that it is possible for an ensemble of uncorrelated repressilators to achieve mutual entrainment through a phase transition, *i.e.* a sudden shift to mutual entrainment as a function of cell-density. Additionally, they have shown that coupling via quorum sensing can also make the ensemble of repressilators robust against stochastic fluctuations.

2.4.3 Coupling Mechanism

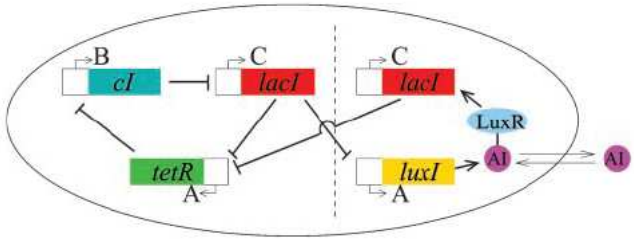

Figure 2.6: Mechanism for coupling repressilators using the auto-inducer (figure from [4])

Cell membranes are usually very selective about letting in or letting out molecules. It is not usually possible for proteins to directly diffuse out of cell membranes and interact with other proteins form other cells. As a result, some new exchange component needs to be introduced into a repressilator cell in order to enable coupling. The diagram above describes the system proposed by Garcia-Ojalvo, Elowitz and Strogatz. They have used the quorum sensing mechanism of the bacterium *Vibrio fischeri*, which is a bioluminicent organism that lives in symbiosis with certain marine hosts forming part of specialised light-emitting organs. The protein LuxI synthesises a small molecule called the auto-inducer (AI). This is the critical coupling component as the AI is the molecule that can diffuse through cell membranes. Another protein LuxR binds with this AI and induces some genes to transcribe creating certain enzymes that ultimately lead to the emission of light. Garcia-Ojalvo *et al.* have placed the protein LacI of the original repressilator under the control of the LuxR-AI

complex as well as placing LuxI under the control of another copy of LacI. The model is described in more detail thus.

- The protein LuxI synthesises the autoinducer AI which can diffuse through cell membranes
- The protein LuxR binds to AI: the LuxR-AI complex induces the expression of a second copy of the gene *lacI*
- The expression of the gene *luxR* is controlled by the original protein LacI
- A second feedback loop appears in the cell
- Increased LacI concentration inhibits the rate of expression of the genes *tetR* and *luxI*
- Decreased LuxI concentration decreases concentration of LuxR-AI complex
- Decreased LuxR-AI concentration decreases *lacI* concentration
- The two LacI transcripts are assumed to be identical

2.4.4 Conclusions

A description of the mathematical modelling as well as the simulation results by Garcia-Ojalvo *et al.* is given in the next chapter. Briefly, what they conclude is that quorum sensing can be an especially effective method of intercell couping between repressilators, and this can lead to global synchronisation among a highly heterogeneous ensemble of repressilators. Also, they found that such an ensemble is also robust to random phase drifts of the individual oscillators due to noise[4]. It is an interesting result and seems to suggest that the problems faced by Elowitz and Leibler[3] using a single repressilator with noise can be resolved by using a large number of repressilators coupled by quorum sensing, even in the presence of noise. The positive effect of coupling is one of the points by which this model differs from the original design of the repressilator, which had only negative feedback, and it is possible that such positive feedback is instrumental in providing mutual entrainment and robustness against stochasticity. However, it is to be remembered (as Garcia-Ojalvo, Elowitz and Strogatz themselves admit

in their paper) that quorum sensing is but one of the many different ways of coupling cells together. Examples given in their paper include the coupling in the sinoatrial node[6] of the heart which is electronic and is mediated by gap junctions. Also, in the suprachiasmatic nuclei[7] of the mammalain circadian pacemaker, coupling is proposed to occur due to the diffusion of an inhibitory neurotransmitter.

Thus such examples of naturally occurring robust clocks cannot be modelled by the mechanism proposed by Garcia-Ojalvo *et al.*. A close analog as mentioned in the paper might be in metabolic synchrony observed in yeast cells. So the proposed system only partially reflects the natural world. However, it is true that the variability of types of intercell coupling in naturally occurring oscillators is too great to allow for an effective model by any single simple mechanism. From that viewpoint, the work done by these three researchers is important. However, according to a private communication from Mukund Thattai, experimentalists have so far been unable to reproduce the simulation results given above. A possible reason for this is that the strength of intercell coupling by quorum sensing might not be strong enough to overcome noise and fluctuation effects.

[6] The sinoatrial node of the heart is what is commonly known as the pacemaker. It is responsible for maintaining sinus rhythm, that is the regular beating of the heart.

[7] The suprachiasmatic nucleus is a bilateral region of the brain located in the hypothalamus responsible for carrying out circadian rhythms within the body.

Chapter 3

Mathematical Modelling

3.1 Introduction

The *dharma* of physicists is to model physical systems (simple or complicated) and the tool that is used to achieve such ends is Mathematics. However, solving mathematical equations can sometimes be impossible due to non-linearity of the equations or the complexity of the problem. Computer simulations then come to the rescue. The repressilator too has been modelled mathematically.[3][6][4] Elowitz and Leibler, in addition to performing experiments, also modelled the repressilator mathematically[3]. They used deterministic modelling using the chemical kinetics approach which involved Michaelis-Menten dynamics that regulate the rate of formation and reactions of enzymes. They also modelled the repressilator stochastically, and used the Gillespie algorithm[2] to solve it. Adiel Loinger and Ofer Biham[6] chose to adopt the rate equation approach for the deterministic case and the master equation approach for the stochastic case. Finally, Garcia-Ojalvo, Elowitz and Strogatz[4] modelled the quorum sensing problem using the Michaelis-Menten approach. They also added a Gaussian noise component of the Ornstein-Uhlenbeck type[9] to model the stochastic case.

3.2 Elowitz and Leibler's model

3.2.1 Deterministic: Michaelis-Menten Kinetics

The dynamic variables in this model are the repressor proteins and the mRNA molecules. As defined earlier in Chapter 2, gene expression is a combination of two processes: transcription, in which a gene produces messenger RNA, and

translation, in which proteins are produced from mRNA. Each of the three proteins/mRNA molecules were considered to be identically behaved. There are six coupled first-order differential equations.

$$\frac{dm_i}{dt} = -m_i + \frac{\alpha}{\left(1 + p_j^n\right)} + \alpha_0 \tag{3.1}$$

$$\frac{dp_i}{dt} = -\beta\left(p_i - m_i\right) \tag{3.2}$$

The indices i and j run from 1 to 3. Here i =*lacI*, *tetR*, *cI*, while j =*cI*,*lacI*,*tetR*. The quantities p_i and m_i are the concentrations of the repressing protein and the mRNA respectively and suitably normalized ('concentration' here means the average copy number[1] per cell). The parameter $\alpha + \alpha_0$ is the rate of production of the protein in absence of the repressing protein. In the presence of a repressor the rate drops to $\frac{\alpha}{\left(1+p_j^n\right)} + \alpha_0$, where the first term gives the effect of the concentration of the repressor protein modified by the Hill coefficient n, and α_0 is the 'leakiness quotient' that describes translation rate independent of the repressor. So the process of gene expression in the repressilator circuit is divided into two parts: the rate part that is modified by the concentration of the repressor and the other part that is not. The parameter β is the ratio of the decay rates of the protein and the mRNA. The normalization of the protein and mRNA concentration are thus: m_i is normalized by the translation efficiency which is the average number of proteins produced per mRNA molecule, while p_i is normalized by the quantity K_m (called the Michaelis constant) which is the number of repressor necessary to half-maximally repress a promoter (*i.e.* an mRNA molecule which performs translation). The Hill coefficient is a measure of the degree of cooperativity of the attaching molecules (here the repressors). A Hill coefficient of 1 indicates that the effect of binding a repressor does not depend on the number of repressors already present. A Hill coefficient greater than one indicates positive cooperativity so that the effect of the binding of each new repressor is enhanced by the number of repressors already bound. A Hill coefficient less than one indicates the the effect of existing repressors decreases the effect of each new binding.

[1] Copy number means the average number of molecules of a gene per genome contained in a cell. Genome is the complete package of genetic material for a living thing. A copy of the genome is found in most cells.

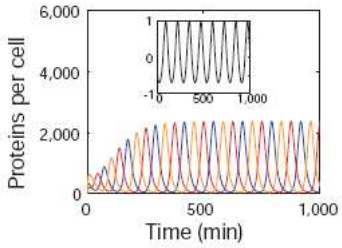

Figure 3.1: Oscillation in the levels of the three repressor proteins in the deterministic case. The inset shows the normalized autocorrelation function of the first protein. The parameter values used by Elowitz and Leibler were as follows: average translation efficiency = 20 proteins per transcript, Hill coefficient $n = 2$, protein half-life = 10 minutes, mRNA half-life = 2 minutes, $K_m = 40$ repressors per cell (figure from [3]).

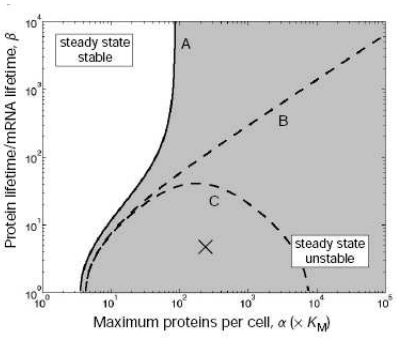

Figure 3.2: Analysis of the stability of the steady state against parameters β and $\alpha \times K_m$. Stable and unstable regions in the β-α parameter space with respect to the steady state are shown. The cross mark in the unstable region of the graph corresponds to the parameter values of figure 3.1. It is in the unstable steady state that oscillations occur (figure from [3]).

The shaded region in figure 3.2 shows the region of parameter space for which oscillations take place.

3.2.2 Stochastic: Gillespie Algorithm

Elowitz and Leibler used the Gillespie SSA algorithm[2] to solve for the stochastic case (following the Gillespie prescription[2]). Parameter values used were chosen keeping in mind that they approximately be similar to those chosen in the deterministic case. The output was the following figure.

It can clearly be seen (figure 3.3) that there are oscillations which persist, but there is large variability. As a result, the autocorrelation time is finite

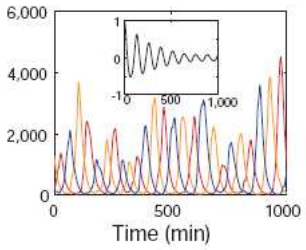

Figure 3.3: Oscillation in the levels of the three repressor proteins in the stochastic case (y-axis is proteins per cell). The inset shows the normalized autocorrelation function of the first protein. A similar set of values of parameters to those of figure 3.1 were used (figure from [3]).

(approximately two periods). So stochasticity seems to be a nuisance for regular oscillations in the single repressilator both experimentally (see Chapter 2) and in simulations.

In this model, oscillations are not seen if the Hill coefficient is taken as 1. this means that a single repressor attached to each transcript carries out the job of repression. This is called non-cooperative binding. As the Hill coefficient is increased to 2, oscillations are observed for suitable values of the parameters only if the mRNA level is included. For Hill coefficient 3 or larger, oscillations take place even if the mRNA level is not included in he analysis. This seems to show that there is positive cooperativity in the system.

3.3 Loinger and Biham's model

Adiel Loinger and Ofer Biham have recently studied[6] the repressilator through computer simulations. In the Elowitz-Leibler model, oscillations were not observed for Hill coefficient $n = 1$, even if the mRNA level was included. However, Loinger and Biham have attempted to check for oscillations without cooperative binding ($n = 1$) using the following reclassifications of proteins and ignoring the mRNA level. They have subdivided the proteins into two types: free and bound. As mentioned earlier, a protein represses the next gene and slows down its rate of expression. Loinger and Biham theorise that in order to bring about its repressive action on the gene next in line, a protein needs to be bound to the gene. If a protein expressed by a gene is not bound to the gene next in line, it is then considered to be 'free' and 'inactive'. If it is bound, then it is

considered to be carrying out repression and is labelled 'bound' or 'active'. The word 'binding' implies that the protein has bound itself to the mRNA next in line and is inhibiting transcription, which effectively represses expression. The absence of cooperative binding also means that there is only one binding site per mRNA molecule (or gene).

The present problem can be studied both deterministically and stochastically. We begin by studying the deterministic case.

3.3.1 Deterministic treatment

The concentration[2] of the i^{th} protein p_i is subdivided into the above-mentioned two classes:

a) Free (or inactive) : p_i^f

b) Bound (or active) : p_i^b

Of course $p_i = p_i^f + p_i^b$, with i going from 1 to 3.

Processes

The processes involved in the time-evolution of the two classes of proteins are thus:

1. Degradation: Both the free and bound proteins degrade at certain rates. The parameter γ_f represents the rate of degradation of the free proteins, while the parameter γ_b indicates the rate of degradation of the bound proteins.

2. Expression: Free proteins are expressed from their parent genes at a constant rate g, while the rate of production is repressed by the previous bound protein. The word 'previous' implies that if it is the rate of production of the i^{th} protein under consideration, then the $(i-1)^{th}$ protein will be called the 'previous' protein. Bound proteins cannot be directly expressed by genes.

3. Exchange of protein classes: Free and bound proteins can convert between themselves. The process of a free protein converting into a bound one is

[2] As defined earlier, 'concentration' here means the average copy number per cell

called 'binding' and the process of a bound proteins converting into a free protein is called 'unbinding'.

a) Binding: Free proteins attach themselves to vacant binding sites and convert to bound proteins. This is the only way bound proteins are formed. The rate of such conversion depends on the number of current free proteins as well as on the number of current vacant binding sites. The latter can be reinterpreted as the number of current bound proteins if the number of binding sites is held constant. A point to be remembered is that a free protein can only bind to the gene (that is, the mRNA) next in line. This means that the i^{th} free protein can only bind to the $(i+1)^{th}$ vacant binding site. Free TetR for example cannot bind itself to *lacI*, but can bind to λ *cI*.

b) Unbinding: Bound proteins can unbind themselves from the mRNA and become free proteins. The rate of such conversion depends only on the number of current bound proteins, as there is no limit to the number of free proteins.

Rate Equations

The above processes give rise to six differential rate equations (two for each of the three proteins). They are (the index i goes over 1, 2 and 3):

$$\frac{dp_i^f}{dt} = g(1 - p_{i-1}^b) - \gamma_f p_i^f - \alpha_0 p_i^f (1 - p_i^b) + \alpha_1 p_i^b \tag{3.3}$$

$$\frac{dp_i^b}{dt} = \alpha_0 p_i^f (1 - p_i^b) - \alpha_1 p_i^b - \gamma_b p_i^b \tag{3.4}$$

Always $p_i = p_i^f + p_i^b$, with i going from 1 to 3.

We look at the two equations term by term.

- Equation (3.3)

  a) First term:

  The first term $g(1 - p_{i-1}^b)$ represents the rate of expression of the i^{th} free protein. As mentioned earlier, g is the constant rate of expression. Here we have normalised the number of binding sites per

cell per protein. Therefore the expression $(1 - p_{i-1}^b)$ is equal to the number of vacant binding sites of the previous $((i-1)^{th})$ protein and indicates the effect of repression by the previous protein. More the value of the $(1 - p_{i-1}^b)$ expression, lesser is the repression on the i^{th} protein.

b) Second term:

The second term $\gamma_f p_i^f$ gives the degradation of the free protein at a rate determined by the parameter γ_f.

c) Third term:

The third term $\alpha_0 p_i^f (1 - p_i^b)$ is a bilinear term. It is the binding term, and indicates the rate of binding of the free protein to a vacant binding site. This term depends on the free protein concentration as well as the number of vacant binding sites of the current (i^{th}) protein. More the value of this expression, greater is the probability of conversion of a free protein into a bound one.

d) Fourth term:

The fourth term $\alpha_1 p_i^b$ is a linear term. It is the unbinding term, and indicates the unbinding of the bound proteins from the binding sites, conversion into free proteins and freeing up of the binding sites. This term only depends on the bound protein concentration as there is no limit on the number of free proteins. More the value of the expression p_i^b, greater is the possibility of unbinding.

- Equation (3.4)

a) First term:

The first term $\alpha_0 p_i^f (1 - p_i^b)$ is the bilinear binding term, as described earlier.

b) Second term:

The second term $\alpha_0 p_i^f (1 - p_i^b)$ is the linear unbinding term, as described earlier.

c) Third term:

The third term $\gamma_b p_i^b$ gives the degradation of the bound protein at a rate determined by the parameter γ_b.

It is seen that adding equations (3.3) and (3.4) gives the rate equations for the whole protein p_i.

$$\frac{dp_i}{dt} = g(1 - p_{i-1}^b) - \gamma_f p_i^f - \gamma_b p_i^b \tag{3.5}$$

Figure 3.4 shows the processes involved and the function of each term in the rate equations 3.3 and 3.4.

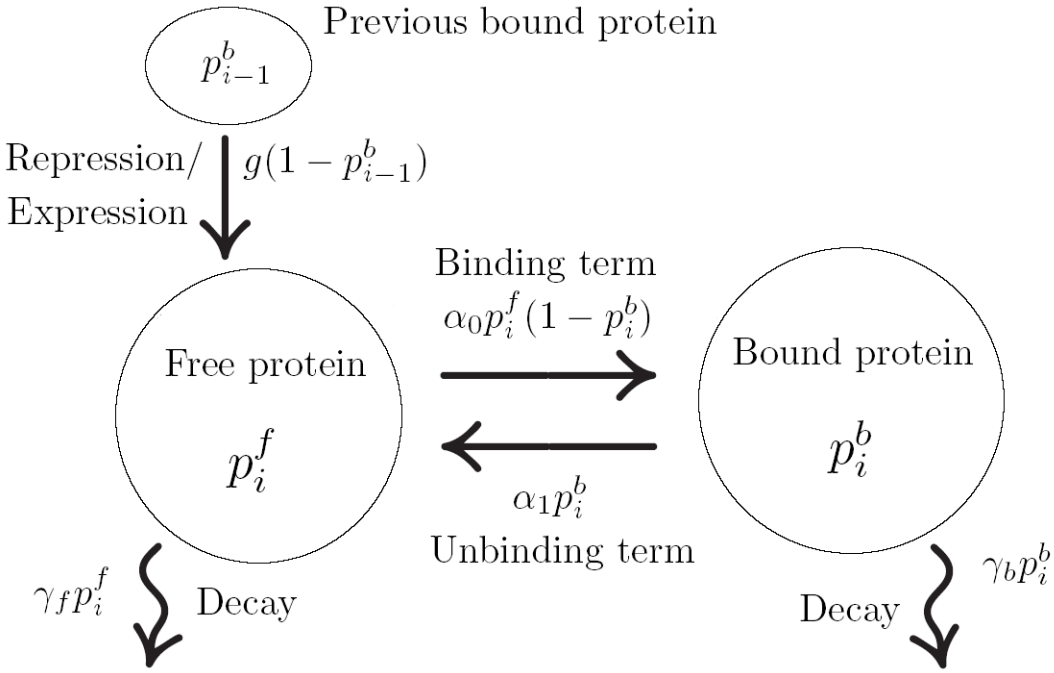

Figure 3.4: Schematic diagram showing the various processes involved and the function of each term in the rate equations 3.3 and 3.4.

3.3.2 Stochastic treatment

The repressors and the binding sites often appear in low copy numbers. As a result, the oscillations are noisy and irregular[7]. Therefore the repressilator

circuit cannot be fully analysed by using the deterministic rate equations approach. It is necessary that we switch to a stochastic approach, and perform our analysis using master equations.

Markov Processes

Master equations are valid for Markov processes[9] only. A Markov process is defined as a stochastic process with the property that for any set of n *successive* times (that is, $t_1 < t_2 < \ldots < t_n$) one has

$$P_{1|n-1}\left(y_n, t_n \mid y_1, t_1; y_2, t_2; \ldots; y_{n-1}, t_{n-1}\right) = P_{1|1}\left(y_n, t_n \mid y_{n-1}, t_{n-1}\right) \quad (3.6)$$

The state n is defined as the state of the system when the variable Y takes the value y_n at time t_n. $P_{1|n-1}$ and $P_{1|1}$ denote conditional probabilities. $P_{1|n-1}$ is the probability density of the variable Y to have taken the value y_n at time t_n, *given* that it had taken the values y_{n-1} at time t_{n-1}, y_{n-2} at time t_{n-2} and so on upto y_1 at time t_1. $P_{1|1}$ is the probability density of the variable Y to have taken the value y_n at time t_n given the value y_{n-1} at time t_{n-1}. $P_{1|1}$ is also called the *transition probability*. The left-hand side of the equation represents the transition probability of a system at the state n given all the previous states in the system's path through phase space. The right hand side denotes the transition probability of the system at state n given the state $n-1$.

The above relation implies that the conditional probability density at t_n, given the value y_{n-1} at t_{n-1}, is uniquely determined and is not affected by any knowledge of the values at earlier times. This means that the evolution of the system at a certain time-step for a discrete Markov process depends only on the value of the system variable at the previous time-step and not on any of the preceding time-steps. In other words, the system has no memory.

Master Equations

Let the range of the system variable Y be a discrete set of states n (as defined earlier). Then we can write the master equation as:

$$\frac{dp_n}{dt} = \sum_{n'} \left\{ W_{nn'}\, p_{n'}(t) - W_{n'n}\, p_n(t) \right\} \quad (3.7)$$

The quantity $W_{nn'}$ is the *transition probability per unit time* from state n' to state n, while the quantity $W_{n'n}$ is the *transition probability per unit time*

from state n to state n'. The quantity $p_n(t)$ is the probability of the system to be in state n at time t, while the quantity $p_{n'}(t)$ is the probability of the system to be in state n' at time t. The LHS dp_n/dt denotes the time evolution of the probability of the system to be in state n at time t. The sum is over all the states $n' \neq n$.

The master equation is thus a gain-loss equation for the probabilities of separate states n. The first term in the equation, $W_{nn'}p_{n'}(t)$, is the gain of state n due to transitions from all other states n'. The second term in the equation, $W_{n'n}p_n(t)$ is the loss due to transitions from state n into other states n'. Also, the transition probability is positive or zero for all cases where the initial and final states are unequal; that is, $W_{nn'} \geq 0$ when $n' \neq n$. Of course, terms with $n = n'$ do not contribute to the sum.

Loinger-Biham master equations

In the stochastic description, we denote the number of the i^{th} free protein by n_i^f, and the number of the i^{th} bound protein by n_i^b. We consider the time evolution of the probability distribution function $P\left(n_1^f, n_2^f, n_3^f, n_1^b, n_2^b, n_3^b\right)$. This is the probability for a cell to include n_i^f copies of the i^{th} free protein and n_i^b copies of the i^{th} bound protein. The quantity n_i^f can take values $0, 1, 2\ldots$, while the quantity n_i^b can only take values $0, 1$, since we have assumed a single binding site. Our master equation will therefore be

$$\begin{aligned}
&\dot{P}(n_1^f, n_2^f, n_3^f, n_1^b, n_2^b, n_3^b) \\
&\quad = \sum_{i=1,2,3} \Big\{ g\left(1 - n_{i-1}^b\right) \left[P\left(\ldots, n_i^f - 1, \ldots, n_1^b, n_2^b, n_3^b\right) - P\left(\vec{n}^f, \vec{n}^b\right)\right] \\
&\qquad + \gamma_f \left[(n_i^f + 1)P\left(\ldots, n_i^f + 1, \ldots, n_1^b, n_2^b, n_3^b\right) - n_i^f P\left(\vec{n}^f, \vec{n}^b\right)\right] \\
&\quad + \alpha_0 \left[(n_i^f + 1)n_i^b P\left(\ldots, n_i^f + 1, \ldots, n_i^b - 1, \ldots\right) - n_i^f (1 - n_i^b) P\left(\vec{n}^f, \vec{n}^b\right)\right] \\
&\qquad + \alpha_1 \left[(1 - n_i^b)P\left(\ldots, n_i^f - 1, \ldots, n_i^b + 1, \ldots\right) - n_i^b P\left(\vec{n}^f, \vec{n}^b\right)\right] \\
&\qquad + \gamma_b \left[(n_i^b + 1)P\left(n_1^f, n_2^f, n_3^f, \ldots, n_i^b + 1, \ldots\right) - n_i^b P\left(\vec{n}^f, \vec{n}^b\right)\right] \Big\} \quad (3.8)
\end{aligned}$$

The expression $\vec{n}^f = (n_1^f, n_2^f, n_3^f)$, while the expression $\vec{n}^b = (n_1^b, n_2^b, n_3^b)$. Therefore, the expression $P\left(\vec{n}^f, \vec{n}^b\right) = P\left(n_1^f, n_2^f, n_3^f, n_1^b, n_2^b, n_3^b\right)$.

Lets look at equation (3.8) term by term.

- LHS:

  This term is the time evolution of the probability distribution function $P\left(n_1^f, n_2^f, n_3^f, n_1^b, n_2^b, n_3^b\right)$ defined earlier.

- RHS: First term:

  This term $\left(g\left(1-n_{i-1}^b\right)\left[P\left(\ldots, n_i^f-1, \ldots, n_1^b, n_2^b, n_3^b\right) - P\left(\vec{n}^f, \vec{n}^b\right)\right]\right)$ represents the change of one i^{th} free protein and is the stochastic equivalent of the rate of expression of the i^{th} free protein. The term $g\left(1-n_{i-1}^b\right)$ $\left[P\left(\ldots, n_i^f-1, \ldots, n_1^b, n_2^b, n_3^b\right)\right]$ implies the transition of the system from a state having one less i^{th} free protein to the current state, while the other term $g\left(1-n_{i-1}^b\right)\left[P\left(\vec{n}^f, \vec{n}^b\right)\right]$ implies the transition of the system from the current state to another state having one less i^{th} free protein. It represents, like equation (3.7), a gain-loss scenario. The expression $g\left(1-n_{i-1}^b\right)$ is the transitions probability and plays the roles assigned to $W_{nn'}$ as well as $W_{n'n}$ in equation (3.7). The expression $P\left(\ldots, n_i^f-1, \ldots, n_1^b, n_2^b, n_3^b\right)$ takes the place of $p_{n'}$ in equation (3.7) while $P\left(\vec{n}^f, \vec{n}^b\right)$ substitutes p_n.

- RHS: Second term:

  This term $\gamma_f\left[(n_i^f+1)P\left(\ldots, n_i^f+1, \ldots, n_1^b, n_2^b, n_3^b\right) - n_i^f P\left(\vec{n}^f, \vec{n}^b\right)\right]$ represents the decay of the free protein. The state $(\ldots, n_i^f+1, \ldots, n_1^b, n_2^b, n_3^b)$ is the initial state while $(\vec{n}^f, \vec{n}^b)$ is the current state. The transition from the initial to the current state is accompanied by the transition probability $\gamma_f(n_i^f+1)$, while the reverse transition is weighted by the transition probability $\gamma_f n_i^f$.

- RHS: Third term:

  This term $\left(\alpha_0\left[(n_i^f+1)n_i^b P\left(\ldots, n_i^f+1, \ldots, n_i^b-1, \ldots\right) - n_i^f(1-n_i^b)P\left(\vec{n}^f, \vec{n}^b\right)\right]\right)$ is equivalent to the unbinding term. The transition from state $(\ldots, n_i^f+1, \ldots, n_i^b-1, \ldots)$ to the current state $(\vec{n}^f, \vec{n}^b)$ means that one free protein is decreased and one bound protein is increased. This is weighted by the transition probability $\alpha_0(n_i^f+1)n_i^b$. The reverse transition is weighted by the transition probability $\alpha_0 n_i^f(1-n_i^b)$.

- RHS: Fourth term:

This term $\left(\alpha_1\left[(1-n_i^b)P\left(\ldots,n_i^f-1,\ldots,n_i^b+1,\ldots\right)-n_i^bP\left(\vec{n}^f,\vec{n}^b\right)\right]\right)$ is equivalent to the binding term. The transitions take place between states $(\ldots,n_i^f-1,\ldots,n_i^b+1,\ldots)$ and $(\vec{n}^f,\vec{n}^b)$ which is the current state. Evidently, the number of free proteins increase by one and the number of bound proteins decrease by one in the 'forward' process, while the opposite happens in the 'reverse' process. The 'forward' or 'gain' process is weighted by $\alpha_1(1-n_i^b)$ while the reverse has the transition probability $\alpha_1 n_i^b$.

- RHS: Fifth term:

  This term $\left(\gamma_b\left[(n_i^b+1)P\left(n_1^f,n_2^f,n_3^f,\ldots,n_i^b+1,\ldots\right)-n_i^bP\left(\vec{n}^f,\vec{n}^b\right)\right]\right)$ represents the decay of the bound protein. The transition is between states $(n_1^f,n_2^f,n_3^f,\ldots,n_i^b+1,\ldots)$ and the current state. Here the number of bound proteins decrease or increase by one for the forward and reverse processes. The weights are $\gamma_b(n_i^b+1)$ and $\gamma_b n_i^b$ respectively.

We have made an analysis of the repressilator using the Loinger-Biham rate equations as well as the master equations. The results are given in the next chapter.

3.4 Quorum sensing model

3.4.1 Model (without stochasticity)

The model used by Garcia-Ojalvo *et al.*[4] (refer to chapter 2) in describing their model of intercell coupling through quorum sensing makes use of Michaelis-Menten kinetics in the same spirit as Elowitz and Leibler. The equations for mRNA have terms representing degradation of mRNA and repression by proteins. The term for the *lacI* gene is also modified by the AI concentration.

mRNA equations

$$\frac{da_i}{dt} = -a_i + \frac{\alpha}{1 + C_i^n} \tag{3.9}$$

$$\frac{db_i}{dt} = -b_i + \frac{\alpha}{1 + A_i^n} \tag{3.10}$$

$$\frac{dc_i}{dt} = -c_i + \frac{\alpha}{1 + B_i^n} + \frac{\kappa S_i}{1 + S_i} \tag{3.11}$$

Here a_i, b_i and c_i are the concentrations in the i^{th} cell of the mRNA transcribed from genes *tetR*, *cI* and *lacI* respectively. The corresponding protein concentrations are given here by A_i, B_i and C_i respectively, keeping in mind that the two copies of LacI are assumed to be identical. The quantity S_i is the concentration of AI. The Hill coefficient n indicates (as usual) the level of cooperative binding. The model is non-dimensionalised by measuring time in units of mRNA lifetime which is assumed equal for all genes. The protein levels as well as the AI levels are normalised by their Michaelis constants (as defined earlier in section 1 of this chapter). The parameter α is the dimensionless transcription rate in absence of repressors (equal to $\alpha + \alpha_0$ from section 1 of this chapter), while κ is the maximal contribution to *lacI* transcription in the presence of saturating amounts of AI.

Protein equations

The equations for the rate of change of protein concentrations are given next.

$$\frac{dA_i}{dt} = \beta\,(a_i - A_i) \tag{3.12}$$

The equations for B_i and C_i are similar with b_i and c_i respectively (for the i^{th} cell). The parameter β is the ratio of protein and mRNA degradation rates (*i.e.* the lifetime of mRNA divided by the lifetime of the protein). The mRNA concentrations here have been rescaled by their translational efficiency which, as defined earlier in section 1, is the average number of proteins produced per mRNA molecule.

AI equations

The equations for AI has terms representing degradation, production and inter-cell diffusion.

$$\frac{dS_i}{dt} = -k_{s0}S_i + k_{s1}A_i - \eta(S_i - S_e) \tag{3.13}$$

The first term on the right is the degradation term with the decay coefficient k_{s0}. The second term is the synthesis term while the third term is the intercell coupling term. The lifetimes of TetR and LuxI are assumed to be equal and as a result the concentration of TetR (A_i) also describes the concentration of LuxI here. The coefficient of the coupling term $\eta = \sigma\mathcal{A}/V_c = \delta/V_c$ where σ represents the permeability of the cell membrane, $\mathcal{A}$ is the surface area of the cell membrane and V_c is the volume of the cell.The three parameters have been non-dimensionalised by time-rescaling. The quantity S_e represents the extracellular concentration of AI. Its dynamical equation is

$$\frac{dS_e}{dt} = -k_{se}S_e + \eta_{ext}\sum_{j=1}^{N}(S_j - S_e) \equiv -k_{se}S_e + k_{diff}(\bar{S} - S_e) \tag{3.14}$$

Here $\eta_{ext} = \delta/V_{ext}$. V_{ext} is the total extracellular volume, and the bar represents averaging over all cells. The diffusion rate is given by $k_{diff} = \eta_{ext}N$ and the degradation rate is given by k_{se}. N is the total number of cells.

In the above model, variations in cell density are ignored the concentration of AI is assumed uniform throughout the experimental substrate.

Order parameter

Finally, an order parameter[3] is defined as a measure of phase transition between unsynchronisation and mutual entrainment.

The quantity $b_i(t)$ is the concentration of mRNA in the i^{th} cell. The average of this over all cells is given by

$$M(t) = (1/N)\sum_{i=0}^{n} b_i(t) \tag{3.15}$$

The order parameter R is defined as the ratio of time-variance of $M(t)$ and the variance of b_i averaged over i. So it is basically the ratio of the time-variance of the cell-average of the signal ($b_i(t)$) and the cell-variance of the time-average of the signal.

$$R = \frac{\langle M^2\rangle - \langle M\rangle^2}{\overline{\langle b_i^2\rangle - \langle b_i\rangle^2}} \tag{3.16}$$

If the signals are completely synchronized, then $b_i(t) = b_j(t) = b(t)\ \forall i,j \Rightarrow M(t) = b(t)$. Since all the signals are equal, $\overline{\langle b(t)\rangle} = \langle \bar{b}(t)\rangle$ (time-averaging and cell-averaging commute), and hence $R = 1$ for complete synchronization. For complete unsynchronisation, the signals are completely uncorrelated and as a result the $\langle M^2\rangle = \langle M\rangle^2 \Rightarrow R = 0$.

If R is plotted against some control parameter, a sudden change in its value from 0 to 1 will indicate phase transition.

The control parameter is given by the quantity Q which is defined as

$$Q = \frac{N\delta/V_{ext}}{k_{se} + N\delta/V_{ext}}$$

where $\delta = \sigma\mathcal{A}$. It is assumed that the extracellular volume $V_{ext} >> V_c$, the individual cellular volume. According to this definition, if $N\delta/V_{ext} << k_{se}$ which is the degradation rate of the extracellular AI, then Q is linearly proportional to the cell density (since the permeability σ, the area $\mathcal{A}$ and the extracellular degradation rate k_{se} can be assumed constant). So Q takes the role of the control parameter against which the order parameter R is plotted.

[3] An order parameter is a useful measure of phase transition. The value of the order parameter varies with some control parameter and changes at a phase transition.

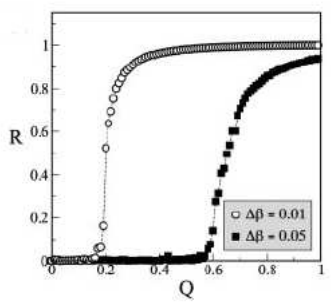

Figure 3.5: Order parameter against Q for the deterministic case.

Results

Simulations carried out by Garcia-Ojalvo, Elowitz and Strogatz yielded the following results. Figure 3.6 shows that synchronization increases as the cell density increases.

Figure 3.5 gives a plot of the order parameter against cell density. It shows that there is a phase transition to synchronization. The parameter $\Delta\beta$ is the spread in parameter values. So Garcia-Ojalvo *et al.* found that there is synchronisation if the number of cells coupling together increases. Also, there is a phase transition to mutual entrainment, which from the figure (f), seems to us to be of the attractive type (as defined in chapter 2) as there is no phase difference between the oscillations.

3.4.2 Model with stochasticity

Extrinsic noise was modelled into the parameter β by adding a Gaussian noise term to it: $\beta \rightarrow \beta_i + \xi_{\mu i}(t)$, where i is the cell-index and μ is the protein-index. The term $\xi_{\mu i}(t)$ was a Gaussian correlated noise of the Ornstein-Uhlenbeck type with zero mean and correlation $\langle \xi_{\mu i}(t)\xi_{\nu j}(t')\rangle = \delta_{\mu\nu}\delta_{ij}\,(D/\tau)\exp(-|t-t'|/\tau)$. The noise is uncorrelated between cells and between genes in each cell. The intensity is D and autocorrelation time is τ. Plots of power spectrums of mRNA concentration as well as order parameter are given in figure 3.7.

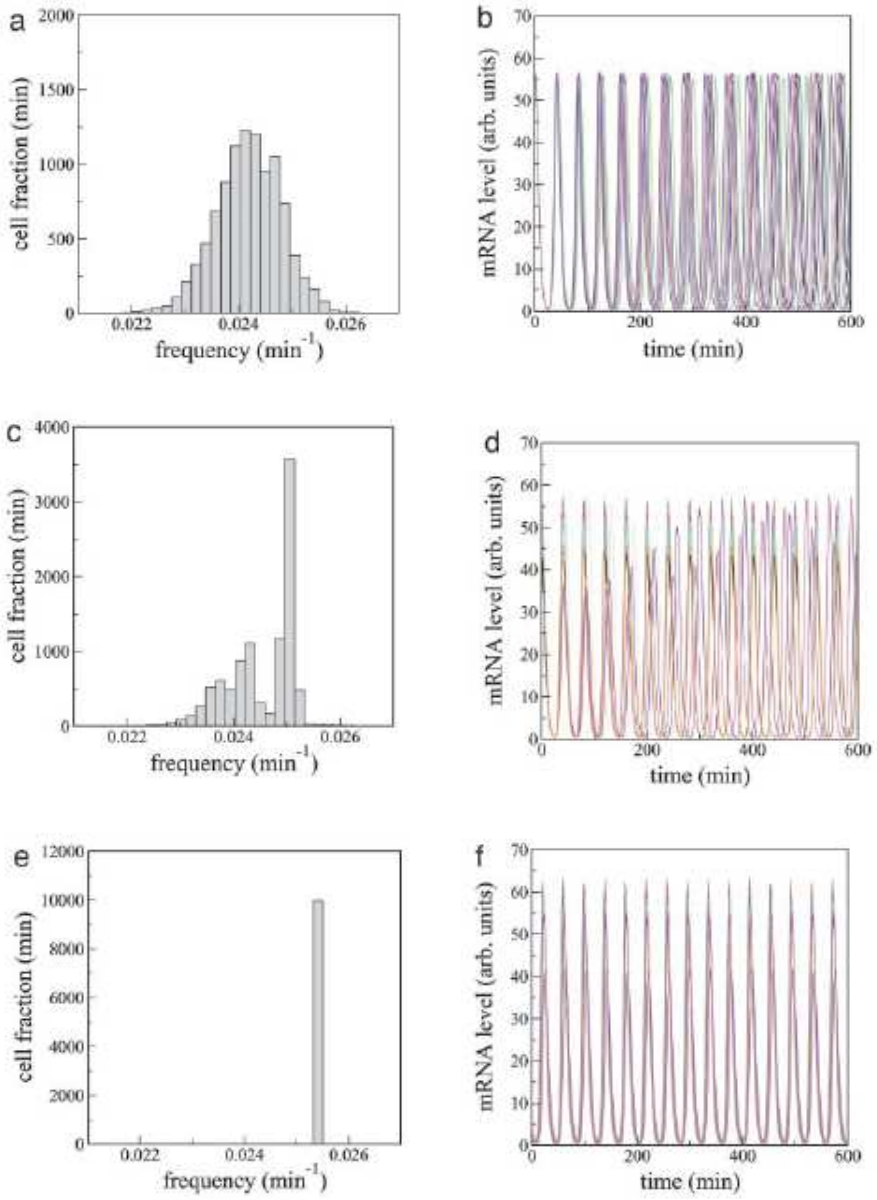

Figure 3.6: Figures in the left column are frequency histograms for different cells. The figures in the right column are signals for different cells plotted against time. Values of Q are 0.4 in the top row, 0.63 in the middle row and 0.8 in the bottom row. The other parameter values used are $\alpha = 216$, $\kappa = 20$, $n = 2.0$, $k_{s0} = 1$, $\eta = 2.0$ and $k_{s1} = 0.01$. The lifetime ratio is chosen from a random Gaussian distribution of mean 1 and standard deviation $\Delta\beta = 0.05$. Figures from[4].

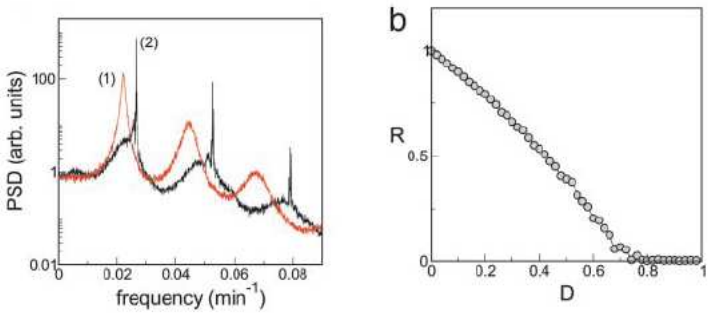

Figure 3.7: Figure on the left gives the power spectrum of the mRNA concentration $b_i(t)$ averaged over 100 repressilators plotted with coupling strength as parameter.$Q = 0$ for curve 1 and $Q = 1$ for curve 2. Noise broadening is seen for the zero-coupling curve (curve 1) while the peaks are far sharper for finite coupling ($Q = 1$ for curve 2). The value of the noise intensity is $D = 0.4$ for both curves (fig. 3.7). Figure on the right gives the order parameter vs noise intensity. $\tau = 15$min, $\eta = 10$ and $\Delta\beta = 0$. The other parameters have the same value as in figure 3.6. Figures from[4].

3.4.3 Conclusion

Garcia-Ojalvo, Elowitz and Strogatz observed[4] that in their stochastic quorum sensing model, synchronisation and coherence is completely lost as the noise intensity increases beyond a certain value ($D \simeq 0.75$ as seen from figure 3.7), no matter the coupling strength. If the noise intensity remains below that critical value then a finite value of the coupling strength can bring about synchronisation. However, they have not studied the effect of internal noise on the oscillations.

Chapter 4

Our Analysis of the Repressilator

4.1 Overview of Work Done

Our study of the repressilator began with the single repressilator. The first job was to non-dimensionalize the rate equations given by Loinger and Biham[6]. Our study was without cooperative binding *i.e.* with the Hill coefficient set to 1. Then we simulated the equations to get concentration vs time graphs for each protein, free or bound. All simulations (including these and subsequent ones) were done in Matlab. After the simulation of the concentration for different values of the remaining parameters, we moved on to simulations of amplitudes, frequencies and waveforms of the protein concentrations against varying parameters. We found definite regions of parameter space for which there were oscillations, as indicated by the above studies of amplitudes and frequencies. The waveforms also differed with parameter values, but were overall closer to sinusoids than relaxation types. These were for the deterministic case. For the stochastic case, we simulated the master equations given by Loinger and Biham using the Gillespie algorithm[2] and obtained oscillation graphs vs time. Finally, we went on to two repressilators coupled by AI and simulated concentrations vs time graphs and phase differences to find out the type of entrainment.

4.2 Single Repressilator: Deterministic

4.2.1 Non-dimensionalization

Why?

One of the problems faced while setting values of parameters to be used in simulations is a baseline to compare them with. The usual *modus operandi* is to non-dimensionalize them against some parameter. This not only reduces the number of parameters in the problem, it also serves a guideline for interpretation of the values of the parameters.

How?

The parameters in the equations are the following (as defined in chapter 3):

- The expression constant g: This is the constant rate of expression of the proteins in the absence of repressors.
- The free-protein degradation constant df (previously γ_f)
- The bound-protein degradation constant db (previously γ_b)
- The binding constant α_0
- The unbinding constant α_1

We non-dimensionalize the Loinger-Biham equations by the parameter g and get the following:

$$\frac{dp_i^f}{d\tau} = (1 - p_{i-1}^b) - \tilde{df} p_i^f - \tilde{\alpha}_0 p_i^f (1 - p_i^b) + \tilde{\alpha}_1 p_i^b \tag{4.1}$$

$$\frac{dp_i^b}{d\tau} = \tilde{\alpha}_0 p_i^f (1 - p_i^b) - \tilde{\alpha}_1 p_i^b - \tilde{db} p_i^b \tag{4.2}$$

where

a) $\tau = tg$

b) $\tilde{df} = df/g$

c) $\tilde{db} = db/g$

d) $\tilde{\alpha}_0 = \alpha_0/g$

e) $\tilde{\alpha}_1 = \alpha_1/g$

4.2.2 Simulation of Concentrations vs Time

We simulated the concentrations of the proteins vs time for different values of the parameters using the RK4[1] algorithm with adaptive time stepping. The time of running the simulation was chosen to have approximately 20 or more stable oscillation cycles. The graphs are given in figure 4.1.

The values of the parameters are the scaled non-dimensionalized 'tilde' values of the parameters (we are removing the tilde forthwith for ease of understanding—now onwards the non-tilde versions will mean the non-dimensionalized parameters). So the system was run with some initial concentration of the first free and bound proteins ($p_1(0) = 1.7$) and zero concentration of the others. [Note: the 'first' protein could be either of the three (TetR, cI or LacI)—the equations are symmetrical with respect to that].

It should be noted that the oscillations, while not relaxational, are not exactly sinusoidal either. This issue will be studied more closely in the section on waveforms.

Further graphs of concentration vs time have also been plotted for varying values of parameters. We do not include them here.

4.2.3 Waveform analysis

We have plotted the concentration of the first bound protein ('first' merely means it has some starting concentration) for varying values of the degradation parameters *df* and *db*. The plots are given in figures 4.2, 4.3 and 4.4. Looking at these figures, we see a clear trend. For low values of *df*, the waveforms are 'top-heavy' with a sort of a plateau at the crests and a spike at the troughs. This is true even for higher values of *db* if *df* be low. For higher values of *df*, the waveforms move closer to a sinusoidal shape.

[1] RK stands for Runge-Kutta

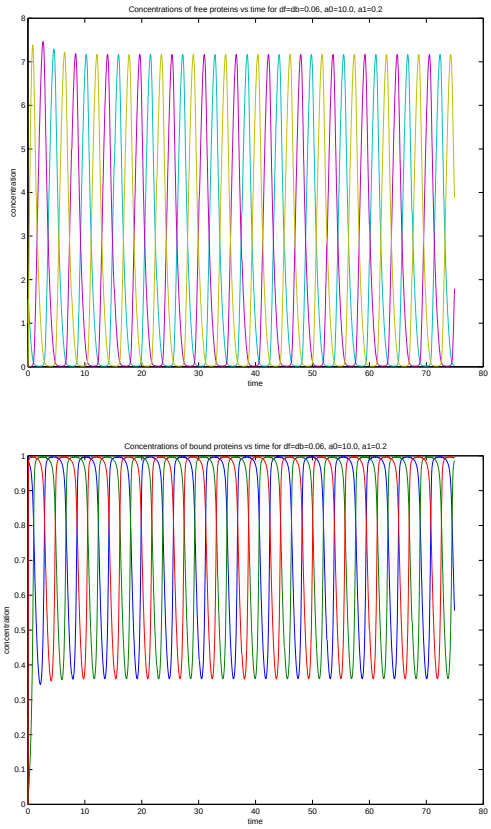

Figure 4.1: The figure on the top gives free protein concentration vs time showing approximately 10 cycles. The figure at the bottom gives bound protein concentration vs time showing approximately 10 cycles. The different colours indicate the different protein concentrations. The parameter values used were $df = 0.06$, $db = 0.06$, $\alpha_0 = 10$, $\alpha_1 = 0.2$. The initial values of the protein concentrations were $p_1^f(0) = 0.7$, $p_i^f(0) = 0.0$ for $i = 2, 3$; $p_1^b(0) = 1.0$, $p_i^b(0) = 0.0$ for $i = 2, 3$.

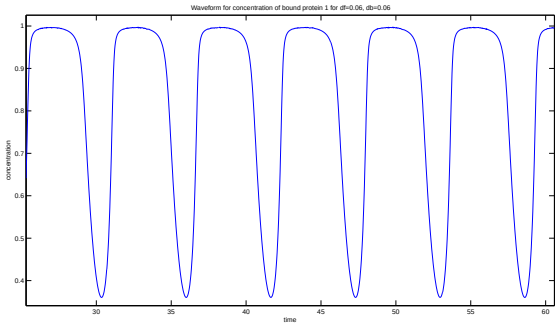

Figure 4.2: Waveform of first bound protein for df=db=0.06

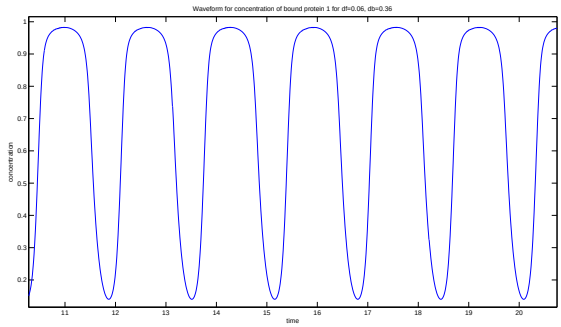

Figure 4.3: Waveform of first bound protein for df=0.06, db=0.36

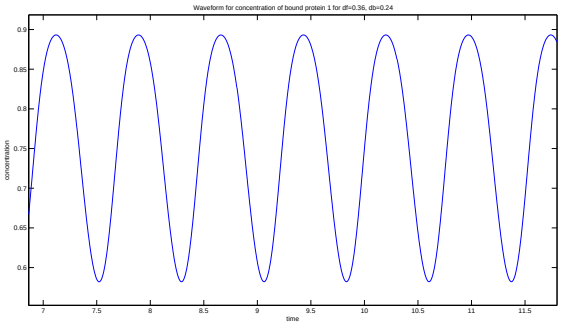

Figure 4.4: Waveform of first bound protein for df=0.36, db=0.24

4.2.4 Amplitude analysis

This analysis was done in the following way. The signal was first detrended[2] and then the absolute value of the signal was taken. Finally a linear fit was done on the signal and the y-intercept was taken. It is quite possible that there are fluctuations at high frequencies but the variations in protein concentration are minimal. This analysis gives an idea of the 'amount' of fluctuations present in the oscillations. We have plotted the intercepts obtained in the way described against the decay parameters *df* and *db*.

Figure 4.5 is a scan of the intercept against *db* for fixed value (0.06) of *df*. It can be seen from this figure that there is substantial amplitude for a large range of *db* (range of *db* in figure is 0.02 to 0.60). However, this is for small $df = 0.06$. We have done a scan for the intercept against both *db* and *df* and 3D plot is

[2]Detrending means setting the mean to zero. This will simply shift the x-axis higher in the graphs.

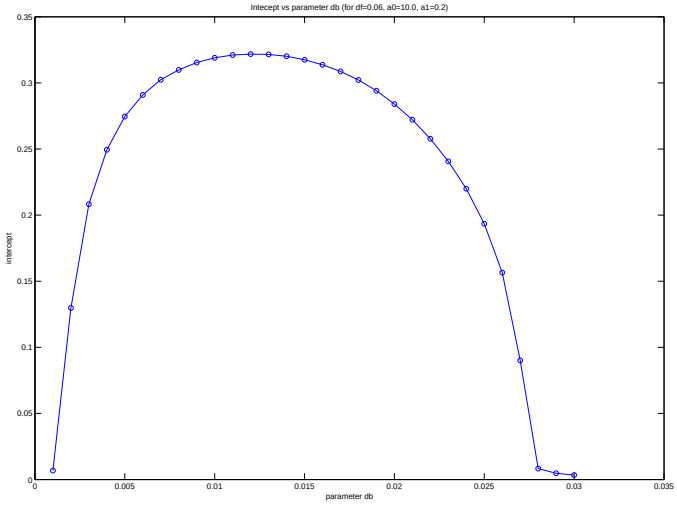

Figure 4.5: Scan of the intercept against the parameter db for $df = 0.06$

given in figure 4.6, and the top-view is given in figure 4.7.

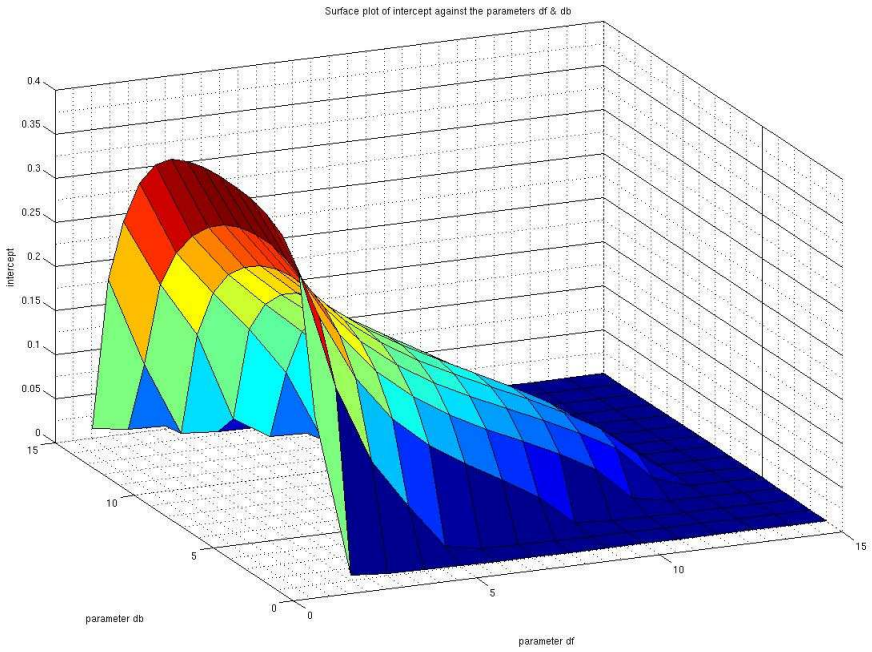

Figure 4.6: 2D scan of the intercept against the parameters df and db. The parameter values are non-scaled values. The ranges of both db and df are 0.02 to 0.58 in steps of 0.04 after scaling by g.

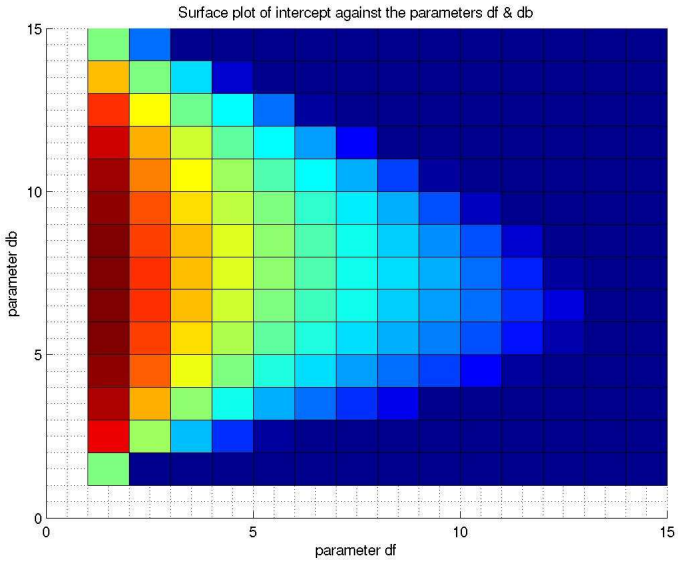

Figure 4.7: Top-view.of the 2D scan of the intercept against the parameters *df* and *db*. The parameter values are non-scaled values. The ranges of both *db* and *df* are 0.02 to 0.58 in steps of 0.04 after scaling by g.

So we see that there are significant oscillations for a certain region of the parameter space of *df* and *db*. For low *df*, there are oscillations almost throughout the range of *db*. But in the higher regime of *df* values, oscillations are found for shorter and shorter ranges of *db* values. At quite high $df > 0.52$, there are no significant oscillations at all.So we conclude that the amplitude of oscillations of the first bound protein are controlled more by the value of the degradation constant of the free proteins than that of the bound proteins.

4.2.5 Frequency analysis

Natural frequencies of oscillations are very important when trying to synchronise oscillators. A study of the frequencies of the oscillations of proteins of a single repressilator is done next. There are two ways of finding out the frequency of a time-dependent signal. The first involves finding the Fourier transform of the signal. The power spectrum of the Fourier transformed signal gives the dominant frequency. This is the frequency obtained from the Fourier transform analysis. Call it ω_{FT}. The other method involves taking the Hilbert transform of the signal. The Hilbert transformed signal is now multiplied by $\imath$ and added

to the original signal to give the 'analytic signal'. The phase of this analytic signal, when differentiated, gives the frequency. Call it ω_{HT}. The final accepted frequency (ω) is the mean of these two frequencies $\omega = (\omega_{FT} + \omega_{HT})/2$. A more detailed description of the Hilbert analysis is given in the Appendix..

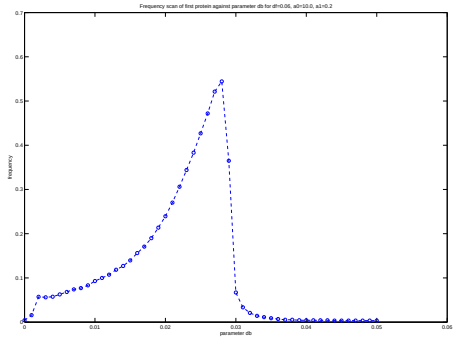

Figure 4.8: Scan of the frequency against the parameter db for $df = 0.06$. The range of db is 0.0 to 1.0 (scaled) for $df = 0.06$.

The interesting thing to observe is that the frequencies seem to rise gradually with increasing db upto 0.56, where it suddenly dips to a very low value. Near the edge of the range the frequency is very close to zero. The following graph (figure (4.9)) is the 3D plot against both df and db. Figure 4.10 is the top-view.

The 3D plot of the frequency scan against df and db (figs 4.9 and 4.9) shows a sharp region where there are finite frequencies. The 2D plot (fig. 4.8) as well as the first 3D plot (fig. 4.9) shows that the frequency increases gradually with db and then falls sharply. This fall happens at lower and lower values of db as the value of df increases. At higher values of df, there is a very small range of db for which there are finite frequencies. This behaviour is qualitatively similar to the behaviour of the amplitudes.

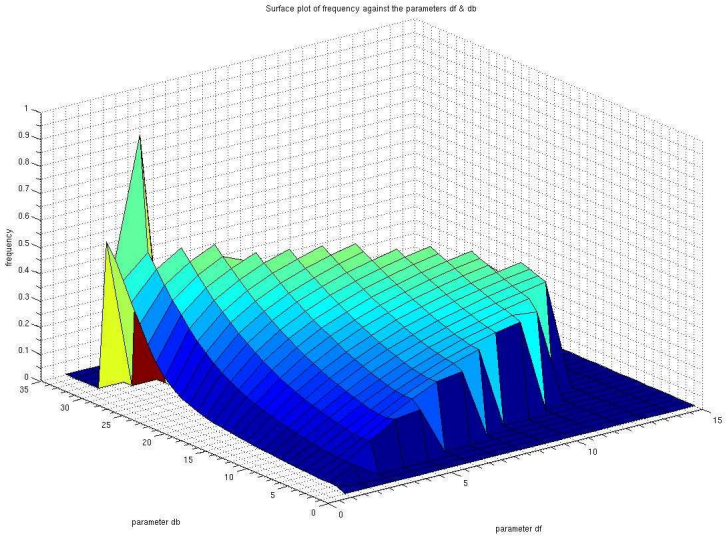

Figure 4.9: 2D scan of the frequency against the parameters *df* and *db*. The range of *db* is 0.02 to 0.70 at intervals of 0.02, while that for *df* is 0.02 to 0.58 at intervals of 0.04.

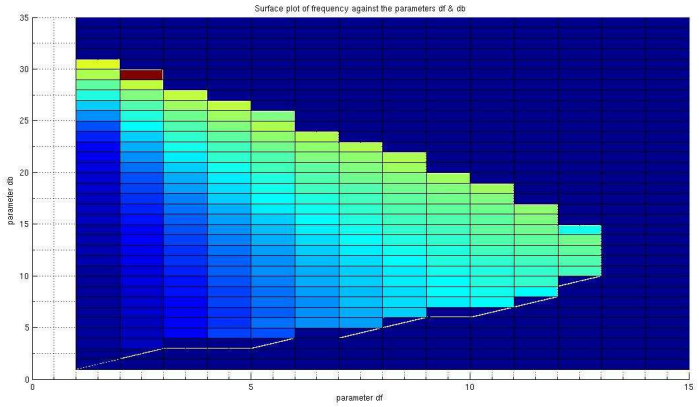

Figure 4.10: 2D scan of the frequency against the parameters *df* and *db*. The range of *db* is 0.02 to 0.70 at intervals of 0.02, while that for *df* is 0.02 to 0.58 at intervals of 0.04.

4.3 Single Repressilator: Stochastic

The stochastic simulations are done by solving the Master Equation (eq. 3.8) using the Gillespie algorithm[2].

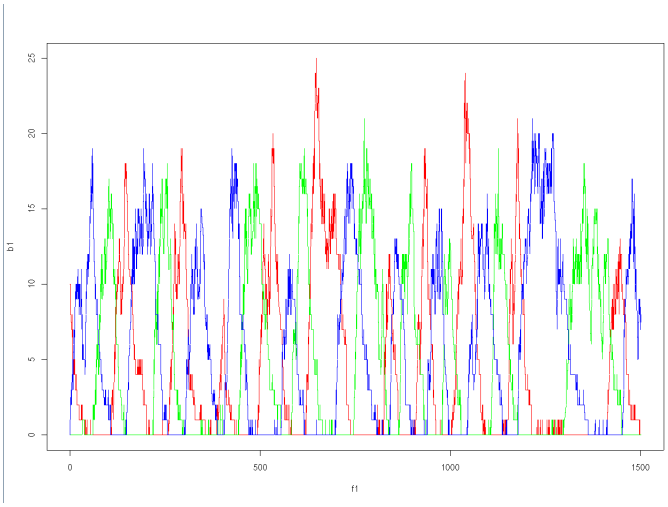

Figure 4.11: Free protein copy number vs time. $\alpha_0 = 10.0$, $\alpha_1 = 0.2$, $df = db = 0.06$, initial first free protein number $f_1(0) = 10.0$, initial first bound protein number $b_1(0) = 1.0$, other initial numbers equalling zero. Time of running is $tmax = 1500$ timesteps. Maximal copy number is $\sim 20 - 25$ with ~ 10 oscillations in 1500 time steps.

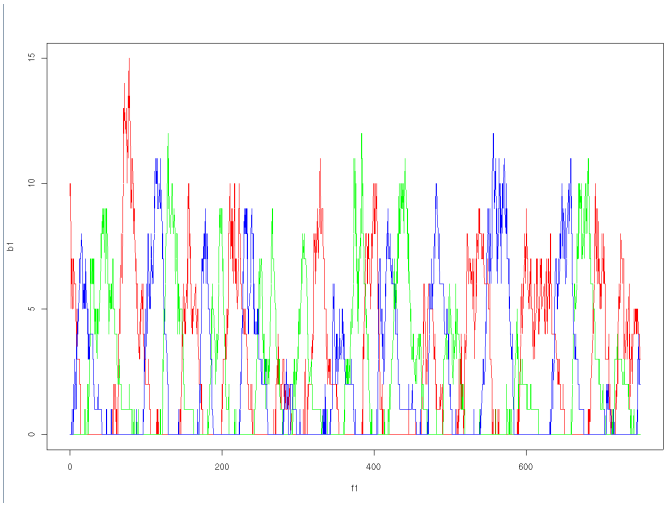

Figure 4.12: Free protein copy number vs time. $\alpha_0 = 10.0$, $\alpha_1 = 0.2$, $df = db = 0.12$, initial first free protein number $f_1(0) = 10.0$, initial first bound protein number $b_1(0) = 1.0$, other initial numbers equalling zero. Time of running is $tmax = 750$ timesteps. Maximal copy number is ~ 10, with $\sim 10 - 15$ oscillations in 750 time-steps.

A comparison of figures 4.11, 4.12, 4.13 and 4.14 exhibit the effect of changing the value of the decay parameters on the frequency and maximal copy number. It is seen that as the values of the decay parameters increase the frequency

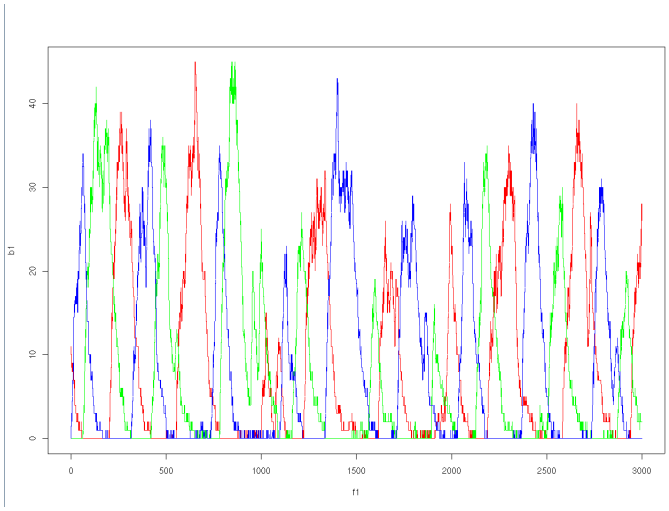

Figure 4.13: Free protein copy number vs time. $\alpha_0 = 10.0$, $\alpha_1 = 0.2$, $df = db = 0.03$, initial first free protein number $f_1(0) = 10.0$, initial first bound protein number $b_1(0) = 1.0$, other initial numbers equalling zero. Time of running is $tmax = 3000$ timesteps. Maximal copy number is $\sim 35 - 40$, with ~ 10 oscillations in 3000 time-steps.

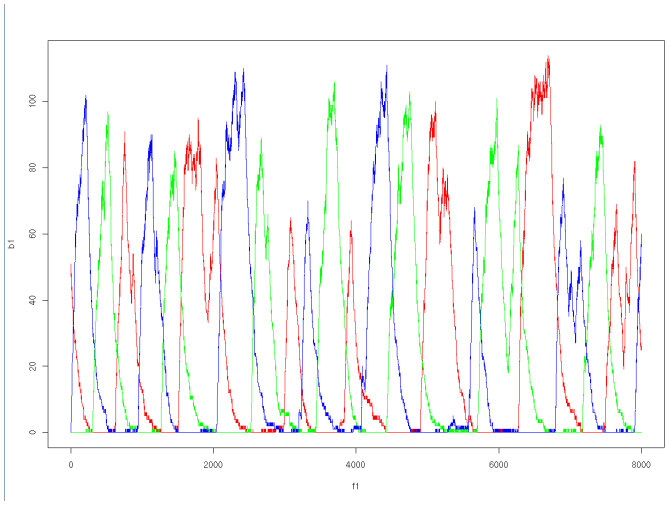

Figure 4.14: Free protein copy number vs time. $\alpha_0 = 10.0$, $\alpha_1 = 0.2$, $df = db = 0.01$, initial first free protein number $f_1(0) = 50.0$, initial first bound protein number $b_1(0) = 1.0$, other initial numbers equalling zero. Time of running is $tmax = 3000$ timesteps. Maximal copy number is ~ 100, with ~ 8 oscillations in 8000 time-steps.

increases but the maximal copy number decreases, and vice versa.

We can summarize the qualitative effects of the parameters thus:

- Increasing decay parameters increases frequencies but decreases maximal copy number.
- Decreasing decay parameters decreases frequencies but increases maximal copy number.
- Changing the initial copy number has no qualitatively discernible effect.
- Changing the binding parameter has no qualitatively discernible effect other than a very slight change in maximal copy number.
- Decreasing the unbinding parameter increases the maximal copy number slightly. However, increasing it decreases the maximal copy number somewhat.

Apart from the effects of changing the decay parameters, the other effects were not very pronounced and in a stochastic system, more rigorous and careful analysis is required to pinpoint the effects. The graphs plotted for examining the change of the binding and unbinding parameters are not given here.

4.4 Two Coupled Repressilators

The coupling is, as described in Chapter 3, done by exchange of the autoinducer AI between cells. This is the quorum sensing mechanism. We have modified the Loinger-Biham rate equations is the following way to introduce the effect of AI coupling.

4.4.1 Modification to Loinger-Biham rate equations

We have two repressilator cells. Our equations for the proteins are

$$\frac{dp_{11}^f}{dt} = g(1-p_{i-1}^a) - \gamma_f p_i^f - \alpha_0 p_i^f(1-p_i^a) + \alpha_1 p_i^a + (p_{11}^f/p_{11}) g_1\, a_1 \qquad (4.3)$$

$$\frac{dp_{12}^a}{dt} = \alpha_0 p_i^f(1-p_i^a) - \alpha_1 p_i^a - \gamma_a p_i^a + (p_{11}^a/p_{11}) g_1\, a_1 \qquad (4.4)$$

Here p^f_{11} is the concentration of the first free protein of the first cell, p^b_{11} is the concentration of the first bound protein of the first cell and p^f_{11} is the total concentration of the first protein of the first cell. The modification terms $(p^f_{11}/p_{11})g_1\,a_1$ and $(p^a_{11}/p_{11})g_1\,a_1$ are parametrized by g_1. These two terms describe the coupling between the second feedback loop to the first (refer to the description in section 2.4.3). So g_1 can be called the *intracell* coupling parameter.

The rate equations of the AI concentrations are given thus:

$$\frac{da_1}{dt} = -g_2 a_1 + g_3 p_{12} - g_4(a_1 - a_2) \tag{4.5}$$

$$\frac{da_2}{dt} = -g_2 a_2 + g_3 p_{22} - g_4(a_2 - a_1) \tag{4.6}$$

Here p_{12} is the second protein in the first cell and p_{22} is the second protein in the second cell. The parameter g_2 is the degradation rate of the AI, g_3 is the induction rate quantified by the concentration of LuxI which is taken to be the same as that of TetR, following the method adopted by Loinger and Biham and described earlier in chapter 3. The *intercell* coupling parameter is g_4 and the terms $-g_4(a_1 - a_2)$ and $-g_4(a_2 - a_1)$ ensure that AI always diffuses out of a cell having higher concentration into the other one.

4.4.2 Entrainment

We need to consider a few cases here.

(A) No intracell or intercell coupling.

In this case we will reproduce the results of section 1 for each cell, and we will not expect the oscillations to be entrained except for the trivial case of same parameters (*i.e.* same natural frequencies) and same initial conditions.

(B) No intercell coupling, but finite intracell coupling

We still do not expect entrainment. However the results, while independent cell-wise, will not be expected to reproduce the single repressilator cases. There is essentially nothing much of interest here. However, the important case is the next one.

(C) Finite intercell and intracell coupling

This will have three sub-cases:

(a) Same initial conditions and parameters: It will a trivial case and the oscillations are expected to perfectly entrain.

(b) Different initial conditions, same parameters: The natural frequencies are same but the initial conditions are different. Still, we expect entrainment as different initial conditions should not affect the phase difference between the cells. however, they might, and it needs to be looked into. The question will be whether the entrainment will be attractive or repulsive.

(c) Different initial conditions and parameters: This is the real deal. The two cells will have different natural frequencies. The study will be carried out using different values of the intercell coupling parameter g_4.

4.4.3 Results

Case (A)

We get the graph (fig. 4.15) keeping initial conditions and parameter values same for both repressilators.

As the concentration graph (fig. 4.15) indicates and the phase curve graph (fig. 4.16) confirms, the two repressilators behave exactly equally. This is what was expected, and this is simply a consistency check. [Note: We have fixed the values of the *df* and *db* parameters at 0.24 each to keep the individual frequencies near the centre of the oscillatory region.] Next we change the initial conditions. We find no entrainment (figs. 4.17 and 4.18) although we do have phase-locking..

Case (B)

Here we introduce finite intracell coupling. We expect no entrainment. We get none. Figures 4.19 and 4.20 attest to that. There is phase-locking. There is not much difference between phase-diagrams 4.18 and 4.20, the former with zero intracell coupling and the latter with low intracell coupling ($g_1 = 0.2$).

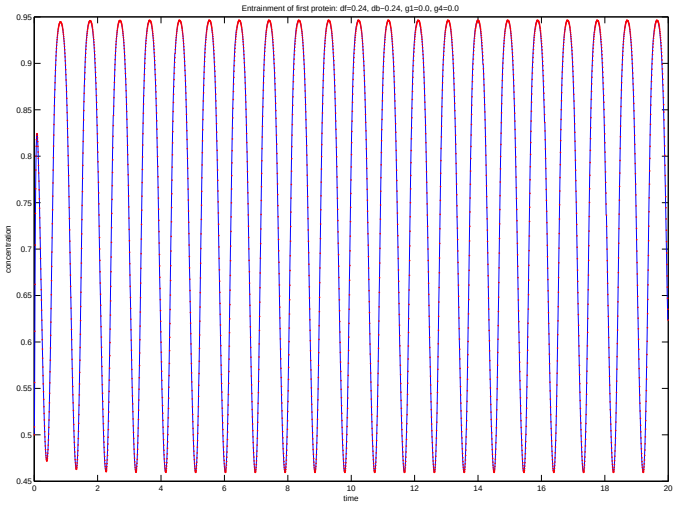

Figure 4.15: First bound protein concentration vs time for two uncoupled repressilators without intracell coupling. The concentrations are indicated by the blue line (for one repressilator) and red dots (for the other). Parameters $\alpha_0 = 10.0$, $\alpha_1 = 0.2$, $df = db = 0.24$, $g_1 = 0.0$, $g_2 = g_3 = 0.2$, $g_4 = 0.0$ for both repressilators. Initial conditions $p^f_{11}(0) = 0.5$, $p^f_{12}(0) = 0.5$.

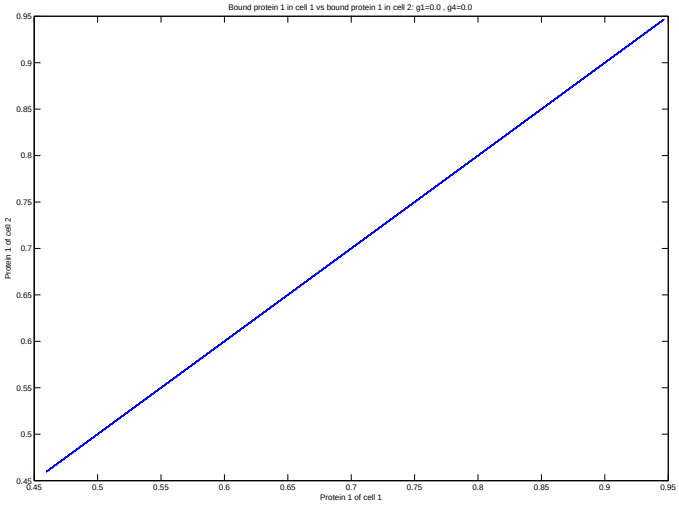

Figure 4.16: Phase curve of the first bound protein concentration of the two uncoupled repressilators without intracell coupling. Parameters $\alpha_0 = 10.0$, $\alpha_1 = 0.2$, $df = db = 0.24$, $g_1 = 0.0$, $g_2 = g_3 = 0.2$, $g_4 = 0.0$ for both repressilators. Initial conditions $p^f_{11}(0) = 0.5$, $p^f_{12}(0) = 0.5$. There is clear entrainment.

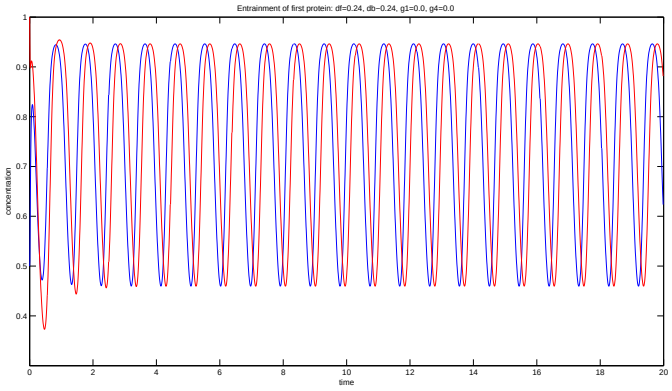

Figure 4.17: First bound protein concentration vs time for two uncoupled repressilators without intracell coupling. Parameters $\alpha_0 = 10.0$, $\alpha_1 = 0.2$, $df = db = 0.24$, $g_1 = 0.0$, $g_2 = g_3 = 0.2$, $g_4 = 0.0$ for both repressilators. Initial conditions $p^f_{11}(0) = 0.5$, $p^f_{12}(0) = 1.0$.

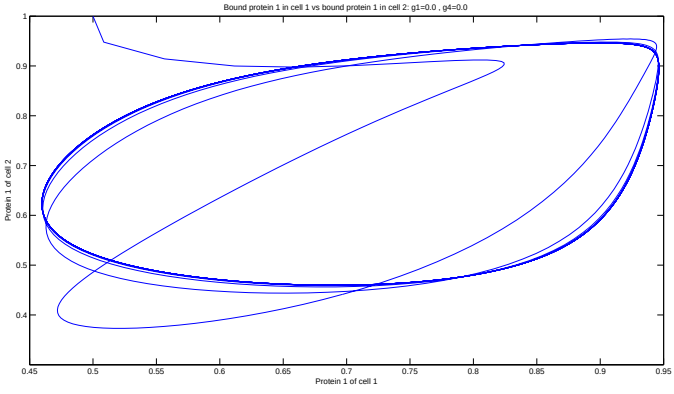

Figure 4.18: Phase curve of the first bound protein concentration of the two uncoupled repressilators without intracell coupling. Parameters $\alpha_0 = 10.0$, $\alpha_1 = 0.2$, $df = db = 0.24$, $g_1 = 0.0$, $g_2 = g_3 = 0.2$, $g_4 = 0.0$ for both repressilators. Initial conditions $p^f_{11}(0) = 0.5$, $p^f_{12}(0) = 1.0$.

Figures 4.21 and 4.22 indicate the case with stronger intracell coupling ($g_1 = 1.4$) keeping all other parameter values same and different initial conditions. It is seen that with increasing strength of the *intracell* coupling, the phase-curve moves towards a smaller limit-cycle. At much higher values of $g_1 (\sim 4.0)$, we get a fixed-point, indicating no oscillations. That is borne out by concentration plots. These plots have been made but are not shown here. Increasing the intracell coupling seems to push the system closer to a non-oscillatory fixed-point. This is

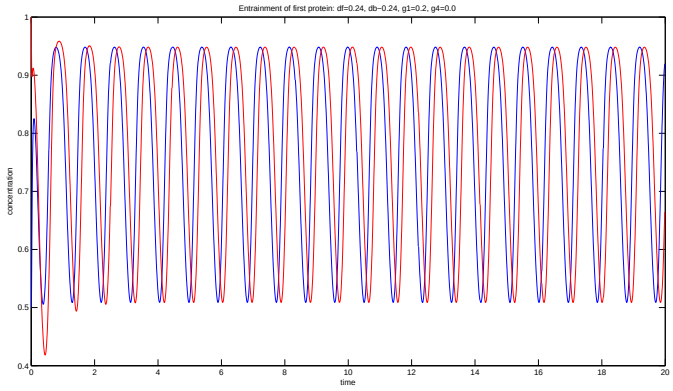

Figure 4.19: First bound protein concentration vs time for two uncoupled repressilators with intracell coupling. Parameters $\alpha_0 = 10.0$, $\alpha_1 = 0.2$, $df = db = 0.24$, $g_1 = 0.2$, $g_2 = g_3 = 0.2$, $g_4 = 0.0$ for both repressilators. Initial conditions $p^f_{11}(0) = 0.5$, $p^f_{12}(0) = 1.0$.

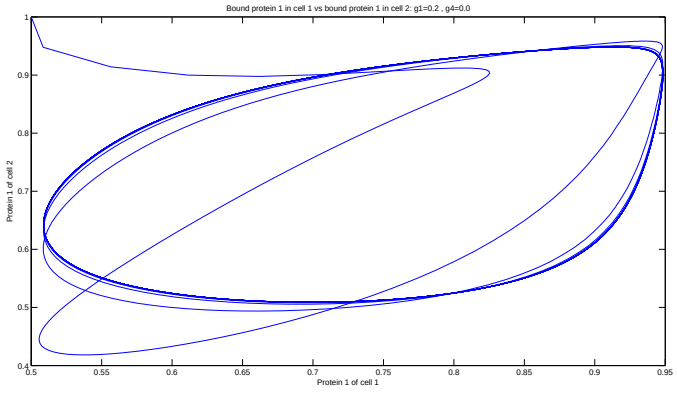

Figure 4.20: Phase curve of the first bound protein concentration of the two uncoupled repressilators without intracell coupling. Parameters $\alpha_0 = 10.0$, $\alpha_1 = 0.2$, $df = db = 0.24$, $g_1 = 0.2$, $g_2 = g_3 = 0.2$, $g_4 = 0.0$ for both repressilators. Initial conditions $p^f_{11}(0) = 0.5$, $p^f_{12}(0) = 1.0$.

because the mechanism underlying the coupling of the second loop (see Chapter 2) is positive feedback which when boosted kills off the oscillations sustained by negative feedback.

A thing to note: although Case (A) with different initial conditions (phase diagram 4.18) and Case (B) with non-zero intracell coupling (phase diagrams 4.20 and 4.22) exhibit no entrainment[3], there *is* phase-locking. We of course

[3]Where entrainment is strictly defined (as earlier) as phase-locking with phase difference

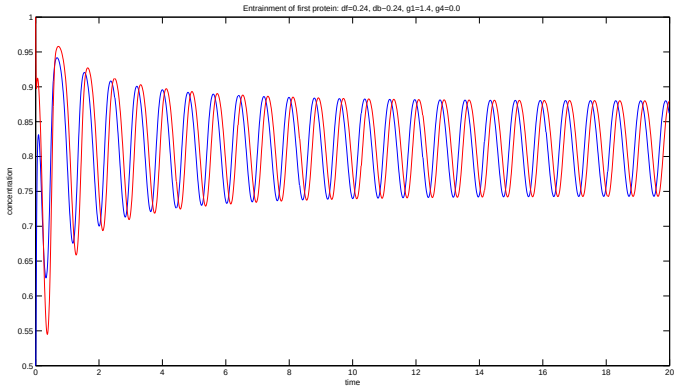

Figure 4.21: First bound protein concentration vs time for two uncoupled repressilators with intracell coupling. Parameters $\alpha_0 = 10.0$, $\alpha_1 = 0.2$, $df = db = 0.24$, $g_1 = 1.4$, $g_2 = g_3 = 0.2$, $g_4 = 0.0$ for both repressilators. Initial conditions $p^f_{11}(0) = 0.5$, $p^f_{12}(0) = 1.0$.

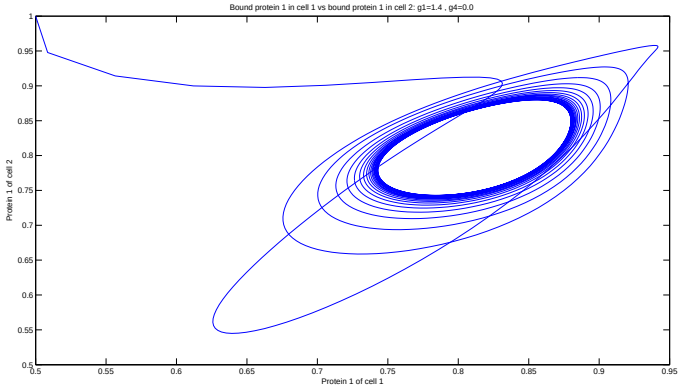

Figure 4.22: Phase curve of the first bound protein concentration of the two uncoupled repressilators with intracell coupling. Parameters $\alpha_0 = 10.0$, $\alpha_1 = 0.2$, $df = db = 0.24$, $g_1 = 1.4$, $g_2 = g_3 = 0.2$, $g_4 = 0.0$ for both repressilators. Initial conditions $p^f_{11}(0) = 0.5$, $p^f_{12}(0) = 1.0$.

do not expect entrainment without intercell coupling. We get phase locking because the natural frequencies of the two repressilators are the same. So we give one set of graphs (figs. 4.23 and 4.24) showing oscillations for Case (B) with *different* frequencies. We change frequencies by tuning the decay parameters. It is seen that there is no longer any limit-cycle type oscillations (fig.4.24).

of 0 or π

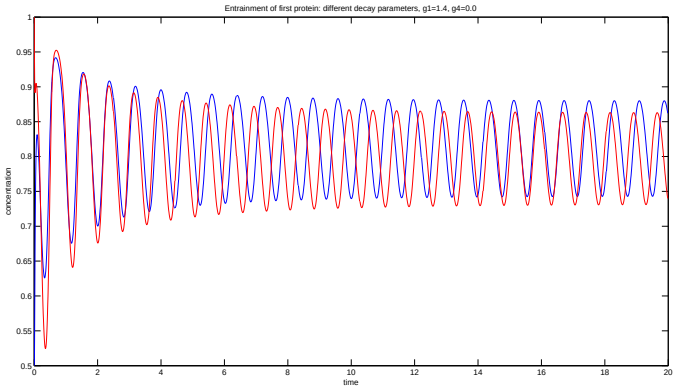

Figure 4.23: First bound protein concentration vs time for two uncoupled repressilators with intracell coupling. Parameters $\alpha_0 = 10.0$, $\alpha_1 = 0.2$, $g_1 = 1.4$, $g_2 = g_3 = 0.2$, $g_4 = 0.0$ for both repressilators. Initial conditions $p^f_{11}(0) = 0.5$, $p^f_{12}(0) = 1.0$. Decay parameters $df = db = 0.24$ for one repressilator, $df = db = 0.264$ for the other.

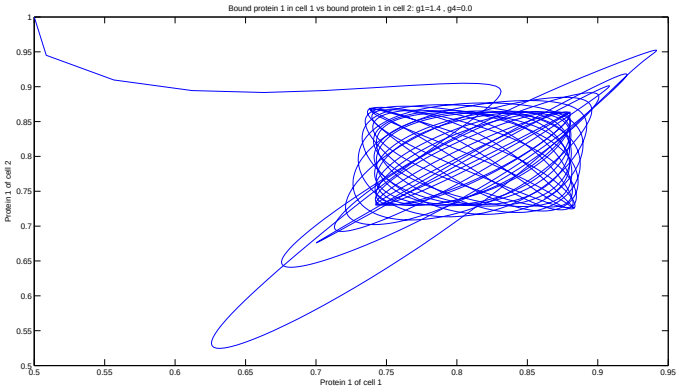

Figure 4.24: Phase curve of the first bound protein concentration of the two uncoupled repressilators with intracell coupling. Parameters $\alpha_0 = 10.0$, $\alpha_1 = 0.2$, $g_1 = 1.4$, $g_2 = g_3 = 0.2$, $g_4 = 0.0$ for both repressilators. Initial conditions $p^f_{11}(0) = 0.5$, $p^f_{12}(0) = 1.0$. Decay parameters $df = db = 0.24$ for one repressilator, $df = db = 0.264$ for the other.

Case (C)

Here we introduce intercell coupling as well. In the first sub-case, we keep natural frequencies and parameters same. We find the trivial case of perfect attractive entrainment (plots not shown). Next we change initial conditions (figs. 4.25 and 4.26).

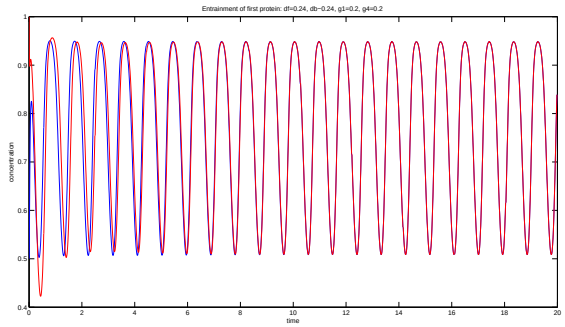

Figure 4.25: First bound protein concentration vs time for two coupled repressilators with intracell coupling. Parameters $\alpha_0 = 10.0$, $\alpha_1 = 0.2$, $df = db = 0.24$, $g_1 = 0.2$, $g_2 = g_3 = 0.2$, $g_4 = 0.2$ for both repressilators. Initial conditions $p^f_{11}(0) = 0.5$, $p^f_{12}(0) = 1.0$.

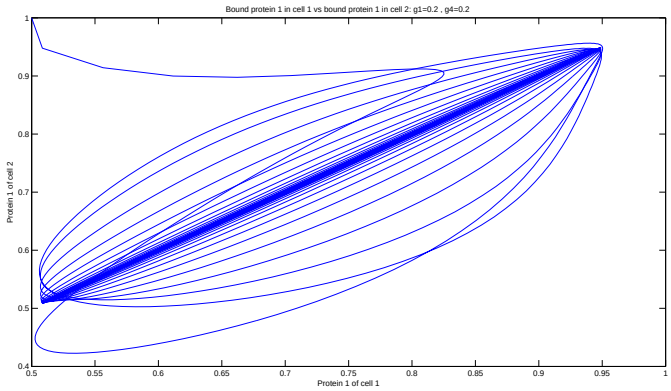

Figure 4.26: Phase curve of the first bound protein of the two coupled repressilators with intracell coupling. Parameters $\alpha_0 = 10.0$, $\alpha_1 = 0.2$, $df = db = 0.24$, $g_1 = 0.2$, $g_2 = g_3 = 0.2$, $g_4 = 0.2$ for both repressilators. Initial conditions $p^f_{11}(0) = 0.5$, $p^f_{12}(0) = 1.0$.

We find from figures 4.25 and 4.26 a transient state after which the two oscillations get entrained attractively. This happens even with coupling strength low for the two repressilators having same natural frequencies.

Finally we come to the case of different initial conditions and natural frequencies. We get figures 4.27, 4.28 and 4.29.

At a high difference of natural frequencies, the oscillations are not even phase-locked (fig. 4.27). However, as the difference in frequencies decreases down from 10% to 1%, there is a tendency to move towards phase-locking (fig.

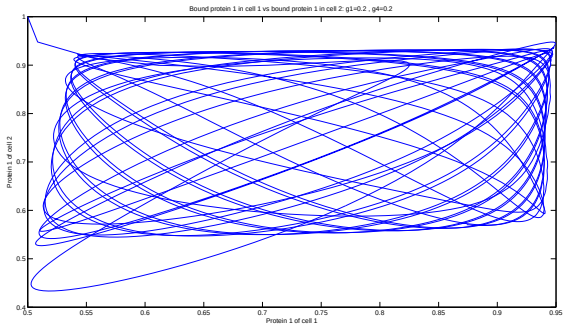

Figure 4.27: Phase curve for 10 % difference in natural frequencies

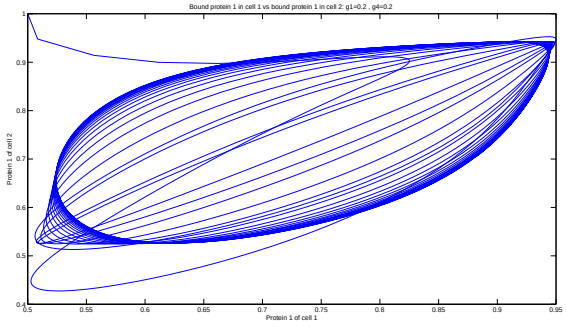

Figure 4.28: Phase curve for 5 % difference in natural frequencies

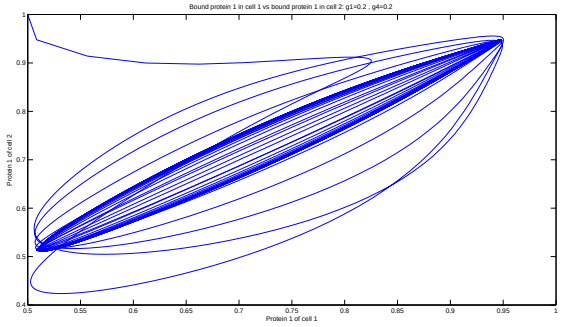

Figure 4.29: Phase curve for 1 % difference in natural frequencies

4.29). However, there is no entrainment as such as could be seen by visual inspection of the phase-curves. Boosting the values of the intracell and/or the intercell parameters also does not affect things greatly, in that even at quite high $g_4 \sim 10.0$, we find no entrainment.

4.4.4 Summary

All this analysis was done for the first bound protein. It seems that there is no entrainment for the two repressilators differing in natural frequencies. If the difference is high, there is no correlation between the phases. However, at lower differences, we get something like phase-locking. However, entrainment still eludes. This was expected.

Our study was on the first bound protein. It is quite possible the other protein oscillations might entrain. That needs be checked. However, it is quite unlikely. Also, our analysis of entrainment was by visual inspection of the phase-curves. This is all right for two repressilators, for the presence or absence of entrainment is quite evident from looking at the pahse curves. However, this approach is not feasible if there are more than two coupled repressilators. In order to quantitatively check for entrainment or phase-locking for more than two repressilators coupled together, we need to define some order parameter.

Chapter 5

Conclusion

The primary motive in studying the repressilator was to try and model natural biological oscillators. The interesting and critical feature of such naturally occurring oscillators was that they preserve their periods of oscillations even if there are externally imposed fluctuations prevalent in natural environments. Examples of such robust oscillators include the circadian oscillator, the pacemaker of the heart etc. Each of these oscillators are made up of collections of cellular oscillators. So there are also internal fluctuations brought about by the different operations of the individual oscillators. In spite of both internal and external fluctuations, biological oscillators are extremely robust and maintain their periodicity. In this thesis, we have reviewed experimental and computer simulation studies of the single repressilator model which showed that single repressilators are not robust to fluctuations. Our review then extended to the coupling of repressilators using the quorum sensing mechanism. Simulations showed that there was robustness against noise for high enough coupling and high enough number of coupled repressilators. However, this has not yet been borne out by experiments.

In our analysis of the repressilator, we started off by studying the single repressilator deterministically using non-dimensionalized rate equations. We managed to isolate regions of parameter space where the single repressilator shows oscillations. Stochastic study by master equations confirmed the result that there are no stable oscillations with stochasticity for the single repressilator. Then we moved on two repressilators coupled using the quorum sensing mechanism. The coupling used between the two repressilators was diffusive in

nature. We studied phase curves of the oscillations of one of the proteins in the two cells. We found as expected that there is attractive entrainment if the natural frequencies are the same, but no entrainment if they are not. The system is phase-locked and comes close to attractive entrainment if the differences in natural frequencies is reduced. These result did not change by increasing coupling strength to large but finite values. However, in general, our conclusion is that at the very least, more than two repressilators need to be coupled in order to get entrainment in the deterministic case when the natural frequencies of the repressilators are different. We have not yet studied stochastic effects on coupling, but it can be expected that the system will move further away from entrainment if stochasticity is included.

Our ultimate aim is to study n of coupled repressilators with stochasticity. We would like to study this in the framework of phase-transitions by defining a suitable order parameter and studying its behaviour with increasing number of cells. We believe that the interplay between stochasticity inherent in a single repressilator, stochasticity from external environmental noise, and the size of the repressilator population will lead to non-trivial structure in the phase-diagram of coupled repressilator circuits.

Appendix

Hilbert Analysis

The aim of Hilbert analysis is to extract the phase of some signal and then maybe the frequency (which is what we have done). The algorithm is as follows. We define the following notations:

- $s(t) \rightarrow$ the signal that is to be analysed
- $s(\omega) \rightarrow$ the fourier transform fo the signal
- $s'(t) \rightarrow$ the Hilbert transform of the signal $s(t)$
- $\tilde{s}(t) \rightarrow$ the 'analytic signal', defined below

Our order of operations is

$$s(t) \xrightarrow[transform]{Hilbert} s'(t) \longrightarrow \tilde{s}(t) = s(t) + is'(t)$$

The Hilbert transform $s'(t)$ of a signal $s(t)$ is defined as

$$s'(t) = \frac{1}{\pi}\mathrm{PV}\int_{-\infty}^{\infty}\frac{s(\tau)}{t-\tau}d\tau \tag{5.1}$$

where PV means principal value.

It is obtained thus:

- Fourier transform[1] the signal: $s(t) \xrightarrow{FFT} s(\omega)$
- Define $s_{>}(\omega) \equiv s(\omega)$ for $\omega > 0$
- Define $s_{<}(\omega) \equiv s(\omega)$ for $\omega < 0$

[1]While coding with Matlab, we have to carry out a Fast Fourier Transform (FFT) due to the discrete nature of the operation of the program

- Shift phase of $s_>(\omega)$ by $-\frac{\pi}{2}$:

  $s_>(\omega) \xrightarrow[-\frac{\pi}{2}]{Phase\ shift} s_>(\omega)\, e^{-\imath\frac{\pi}{2}} \equiv s'_>(\omega)$

- Shift phase of $s_<(\omega)$ by $\frac{\pi}{2}$:

  $s_<(\omega) \xrightarrow[\frac{\pi}{2}]{Phase\ shift} s_<(\omega)\, e^{\imath\frac{\pi}{2}} \equiv s'_<(\omega)$

- Add: $s'(\omega) \equiv s'_>(\omega) + s'_<(\omega)$

- Carry out an inverse fourier transform: $s'(\omega) \xrightarrow{IFFT} s'(t)$, which is the Hilbert transform

We obtain the phase by putting $A(t)e^{\imath\phi(t)} = \tilde{s}(t)$, where we call $A(t)$ the time-dependent amplitude and $\phi(t)$ the time-dependent phase of the analytic signal. The phase of the signal will then be simply $\phi(t)$.

For example, if $s(t) = \cos\omega t$, then $s'(t) = \sin\omega t$, and

$$\tilde{s}(t) = \cos\omega t + \imath \sin\omega t = e^{\imath\omega t} = Ae^{\imath\phi(t)}$$

$$\Rightarrow \phi(t) = \omega t$$

and $A(t) = 1$.

This is consistent, for obviously the phase of the signal $\cos\omega t$ is ωt and amplitude is 1.

Bibliography

[1] N. Barakai and S. Leibler, *Circadian Clocks Limited by Noise*, Nature, 403 (1999), pp. 267–268.

[2] D.T.Gillespie, *Exact Stochastic Simulation of Coupled Chemical Reactions*, Journal of Physical Chemistry, 81 (1977), pp. 2340–2361.

[3] M. B. Elowitz and S. Leibler, *A Synthetic Oscillatory Network of Transcriptional Regulators*, Nature, 403 (2000), pp. 335–338.

[4] J. Garcia-Ojalvo, M. B. Elowitz, and S. H. Strogatz, *Modeling a Synthetic Multicellular Clock: Repressilators Coupled by Quorum Sensing*, PNAS, 101 (2004), pp. 10955–10960.

[5] Y. Kuramoto, *Chemical Oscillations, Waves and Turbulence*, Dover, 2003.

[6] A. Loinger and O. Biham, *Stochastic Simulations of the Repressilator Circuit*, Physical Review E, 76 (2007), p. 051917.

[7] H. H. McAdams and A. Arkin, *Its a Noisy Business! Genetic Regulation at the Nanomolar Scale*, Trends in Genetics, 15 (1999), pp. 65–69.

[8] A. Nandi, S. G., R. Singh, and R. Ramaswamy, *Effective Mechanisms for the Synchronization of Stochastic Oscillators*, Physical Review E, 76 (2007), p. 041136.

[9] N.G. Van Kampen, *Stochastic Processes in Physics and Chemistry*, Elsevier, 3 ed., 2007.

[10] A. Pikovsky, M. Rosenblum, and J. Kürths, *Synchronization: A Universal Cconcept in Nonlinear Sciences*, Cambridge University Press, 2001.

[11] D. Sobel, *Longitude: The True Story of a Lone Genius Who Solved the Greatest Scientific Problem of His Time*, Penguin, 1995.

[12] S. H. Strogatz and I. Stewart, *Coupled Oscillators and Biological Synchronisation*, Scientific American, 269 (1993), pp. 102–109.