Biomolecular Feedback Systems

Domitilla Del Vecchio
MIT

Richard M. Murray
Caltech

Classroom Copy v0.6c, July 11, 2012

Chapter 6
Biological Circuit Components

In this chapter, we describe some simple circuits components that have been constructed in *E. coli* cells using the technology of synthetic biology. We will analyze their behavior employing mainly the tools from Chapter 3 and some of the tools from Chapter 4. The basic knowledge of Chapter 2 will be assumed.

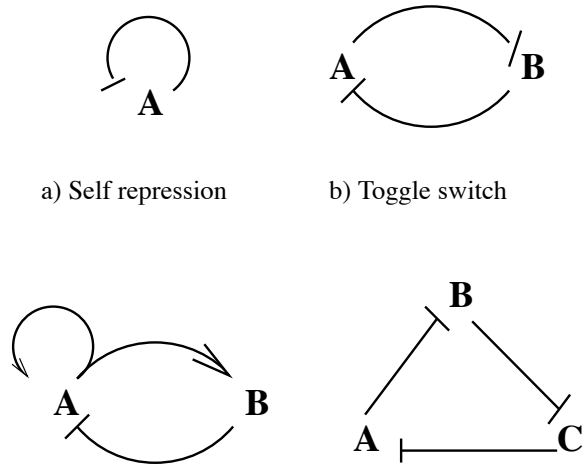
6.1 Introduction to Biological Circuit Design

In Chapter 2 we have introduced a number of core processes and their modeling. These include gene expression, transcriptional regulation, post-translational regulation such as covalent modification of proteins, allosteric regulation of enzymes, activity regulation of transcription factors through inducers, etc. These core processes provide a rich set of functional building blocks, which can be combined together to create circuits with prescribed functionalities.

For example, if we want to create an inverter, a device that returns high output when the input is low and vice versa, we can use a gene regulated by a transcription repressor. If we want to create a signal amplifier, we can employ a cascade of covalent modification cycles. Specifically, if we want the amplifier to be linear, we should tune the amounts of protein substrates to be in smaller values than the Michaelis-Menten constants. If instead we are looking for an almost digital response, we could employ a covalent modification cycle with high amounts of substrates compared to the Michaelis-Menten constants. Furthermore, if we are looking for a fast input/output response, phosphorylation cycles are better candidates than transcriptional systems.

In this chapter and in the next one, we illustrate how one can build circuits with prescribed functionality using some of the building blocks of Chapter 2 and the design techniques illustrated in Chapter 3. We will focus on two types of circuits: gene circuits and signal transduction circuits. In some cases, we will illustrate designs that incorporate both.

A gene circuit is usually depicted by a set of nodes, each representing a gene, connected by unidirectional edges, representing a transcriptional activation or a repression. Inducers will often appear as additional nodes, which activate or inhibit a specific edge. Early examples of such circuits include an activator-repressor system that can display toggle switch or clock behavior [5], a loop oscillator called the repressilator obtained by connecting three inverters in a ring topology [27], a



a) Self repression b) Toggle switch

c) Activator-repressor clock d) Repressilator

Figure 6.1: Early transcriptional circuits that have been fabricated in bacteria *E. coli*: the negatively autoregulated gene [10], the toggle switch [31], the activator-repressor clock [5], and the repressilator [27].

toggle switch obtained connecting two inverters in a ring fashion [31], and an autorepressed circuit [10] (Figure 6.1). Each node represents a gene and each arrow from node Z to node X indicates that the transcription factor encoded in gene z, denoted Z, regulates gene x [3]. If z represses the expression of x, the interaction is represented by Z⊣X. If z activates the expression of x, the interaction is represented by Z→X [3].

Basic synthetic biology technology

Simple synthetic gene circuits can be constituted from a set of (connected) transcriptional components, which are made up by the DNA base-pair sequences that compose the desired promoters, ribosome binding sites, gene coding region, and terminators. We can choose these components from a library of basic interchangeable parts, which are classified based on biochemical properties such as affinity (of promoter, operator, or ribosome binding sites), strength (of a promoter), and efficiency (of a terminator).

The desired sequence of parts is usually assembled on plasmids, which are circular pieces of DNA, separate from the host cell chromosome, with their own origin of replication. These plasmids are then inserted, through a process called transformation in bacteria and transfection in yeast, in the host cell. Once in the host cell, they express the proteins they code for by using the transcription and translation machinery of the cell. There are three main types of plasmids: low copy number (5-10 copies), medium copy number (15-20 copies), and high copy number (up to

hundreds). The copy number reflects the average number of copies of the plasmid inside the host cell. The higher the copy number, the more efficient the plasmid is at replicating itself. The exact number of plasmids in each cell fluctuates stochastically and cannot be exactly controlled.

In order to measure the amounts of proteins of interest, we make use of *reporter genes*. A reporter gene codes for a protein that fluoresces in a specific color (red, blue, green, yellow, etc.) when it is exposed to light of the correct wave-length. For instance, green fluorescent protein (GFP) is a protein with the property that it fluoresces in green when exposed to UV light. It is produced by the jellyfish *Aequoria victoria*, and its gene has been isolated so that it can be used as a reporter. Other fluorescent proteins, such as yellow fluorescent protein (YFP) and red fluorescent protein (RFP) are genetic variations of GFP.

A reporter gene is usually inserted downstream of the gene expressing the protein whose concentration we want to measure. In this case, both genes are under the control of the same promoter and are transcribed into a single mRNA molecule. The mRNA is then translated to protein and the two proteins will be fused together. This technique sometimes affects the functionality of the protein of interest because some of the regulatory sites may be occluded by the fluorescent protein. To prevent this, another viable technique is to clone after the protein of interest the reporter gene under the control of a copy of the same promoter that also controls the expression of the protein. This way the protein is not fused to the reporter protein, which guarantees that the protein function is not affected. Also, the expression levels of both proteins should be close to each other since they are controlled by (different copies of) the same promoter.

Just as fluorescent proteins can be used as a read out of a circuit, inducers function as external inputs that can be used to probe the system. Inducers function by either disabling repressor proteins (negative inducers) or by enabling activator proteins (positive inducers). Two commonly used negative inducers are IPTG and aTc. Isopropyl-β-D-1-thiogalactopyranoside (IPTG) induces activity of beta-galactosidase, which is an enzyme that promotes lactose utilization, through binding and inhibiting the *lac* repressor LacI. The anhydrotetracycline (aTc) binds the wild-type repressor (TetR) and prevents it from binding the Tet operator. Two common positive inducers are arabinose and AHL. Arabinose activates the transcriptional activator AraC, which activates the pBAD promoter. Similarly, AHL is a signaling molecule that activates the LuxR transcription factor. which activates the pLux promoter.

Protein dynamics can be usually altered by the addition of a degradation tag at the end of the coding region. A degradation tag is a sequence of base pairs that adds an amino acid sequence to the functional protein that is recognized by proteases. Proteases then bind to the protein, degrading it into a non-functional molecule. As a consequence, the half life of the protein decreases, resulting into an increased decay rate. Degradation tags are often employed to obtain a faster response of the

protein concentration to input stimulation and to prevent protein accumulation.

6.2 Negative Autoregulation

In this section, we analyze the negatively autoregulated gene of Figure 6.1 and focus on analyzing how the presence of the negative feedback affects the dynamics of the system and how the negative feedback affects the noise properties of the system. This system was introduced in Example 3.6.

Let A denote the concentration of protein A and let A be a transcriptional repressor for its own production. Assuming that the mRNA dynamics are at the quasi-steady state, the ODE model describing the self repressed system is given by

$$\frac{dA}{dt} = \frac{\beta}{1+(A/K)^n} - \delta A.$$

We seek to compare the behavior of this autoregulated system to the behavior of the unregulated one:

$$\frac{dA}{dt} = \beta_0 - \delta A,$$

in which β_0 is the unrepressed production rate.

Dynamic effects of negative autoregulation

As we showed via simulation in Example 2.3, negative autoregulation speeds up the response to perturbations. Hence, the time the system takes to reach its steady state decreases with negative feedback. In this section, we show this result analytically by employing linearization about the steady state and by explicitly calculating the time the system takes to reach it.

Let $A_e = \beta_0/\delta$ be the steady state of the unregulated system and let $z = A - A_e$ denote the perturbation with respect to such a steady state. The dynamics of z are given by

$$\frac{dz}{dt} = -\delta z.$$

Given a small initial perturbation z_0, the time response of z is given by the exponential

$$z(t) = z_0 e^{-\delta t}.$$

The "half-life" of the signal $z(t)$ is the time it takes to reach half of z_0. This is a common measure for the speed of response of a system to an initial perturbation. Simple mathematical calculation shows that $t_{\text{half}} = \ln(2)/\delta$.

Let now A_e be the steady state of the autoregulated system. Assuming that the perturbation z with respect to such a steady state is small enough, we can employ

linearization to describe the dynamics of z. These dynamics are given by

$$\frac{dz}{dt} = -\delta z,$$

in which

$$\delta = \delta + \frac{nA_e^{n-1}/K^n}{(1+(A_e/K)^n)^2}.$$

In this case, we have that $t_{\text{half}} = \ln(2)/\delta$.

Since $\delta > \delta$ (independently of the steady state A_e), we have that the dynamic response to a perturbation is faster in the system with negative autoregulation. This confirms the simulation findings of Example 2.3.

Noise filtering

In this section, we investigate the effect of the negative feedback on the noise spectrum of the system. In order to do this, we employ the Langevin modeling framework and determine the frequency response to the noise on the various reaction channels. We perform two different studies. In the first one, we assume that the decay rate of the protein is much smaller than that of the mRNA. As a consequence, the mRNA is at its quasi-steady state and we focus on the dynamics of the protein only. In the second study, we investigate the consequence of having the mRNA and protein decay rates in the same range so that the quasi-steady state assumption cannot be made. In either case, we study both the open loop system and the closed loop system (the system with negative autoregulation) and compare the corresponding frequency responses.

Assuming mRNA at the quasi-steady state

In this case, the reactions for the open loop system are given by

$$\mathrm{R}_1\colon \mathrm{p} \xrightarrow{\beta_0} \mathrm{A} + \mathrm{p}, \qquad \mathrm{R}_2\colon \mathrm{A} \xrightarrow{\delta} \emptyset,$$

in which β_0 is the constitutive production rate, p is the DNA promoter, and δ is the decay rate of the protein. Since the concentration of DNA promoter p is not changed by these reactions, it is a constant, which we call p_{tot}.

Employing the Langevin equation (4.9) of Section 4.1 and letting n_A denote the real-valued number of molecules of A and by n_p the real-valued number of molecules of p, we obtain

$$\frac{dn_A}{dt} = \beta_0 n_p - \delta n_A + \sqrt{\beta_0 n_p} N_1 - \sqrt{\delta n_A} N_2,$$

in which N_1 and N_2 are the noises on the production reaction and on the decay reaction, respectively. By denoting $A = n_A/\Omega$ the concentration of A and $p = n_p/\Omega = p_{\text{tot}}$

the concentration of p, we have that

$$\frac{dA}{dt} = \beta_0 p_{\text{tot}} - \delta A + \frac{1}{\sqrt{\Omega}}(\sqrt{\beta_0 p_{\text{tot}}} N_1 - \sqrt{\delta A} N_2).$$

This is a linear system and therefore we can calculate the frequency response to any of the two inputs N_1 and N_2. The frequency response to input N_1 is given by

$$G_{AN_1}(\omega) = \frac{\sqrt{\beta_0 p_{\text{tot}}/\Omega}}{\sqrt{\omega^2 + \delta^2}}.$$

We now consider the autoregulated system. The reactions are given by

$$\mathrm{R}_1\colon \mathrm{p} \xrightarrow{\beta} \mathrm{A} + \mathrm{p}, \qquad \mathrm{R}_2\colon \mathrm{A} \xrightarrow{\delta} \emptyset,$$

$$\mathrm{R}_3\colon \mathrm{A} + \mathrm{p} \xrightarrow{a} \mathrm{C}, \qquad \mathrm{R}_4\colon \mathrm{C} \xrightarrow{d} \mathrm{A} + \mathrm{p}, \qquad p_{\text{tot}} = p + C.$$

Employing the Langevin equation (4.9) of Section 4.1 and dividing both sides of the equation to obtain concentrations, we obtain

$$\begin{aligned}
\frac{dp}{dt} &= -aAp + d(p_{\text{tot}} - p) + \frac{1}{\sqrt{\Omega}}(-\sqrt{aAp} N_3 + \\
&\quad \sqrt{d(p_{\text{tot}} - p)} N_4) \\
\frac{dA}{dt} &= \beta p - \delta A - aAp + d(p_{\text{tot}} - p) + \frac{1}{\sqrt{\Omega}}(\sqrt{\beta p} N_1 - \sqrt{\delta A} N_2 - \sqrt{aAp} N_3 + \\
&\quad \sqrt{d(p_{\text{tot}} - p)} N_4),
\end{aligned}$$

in which N_3 and N_4 are the noises on the association and o the dissociation reactions, respectively. Letting $K_{\text{d}} = d/a, \Gamma_1(t) = \frac{1}{\sqrt{\Omega}}(-\sqrt{aAp/K_{\text{d}}} N_3 + \sqrt{d(p_{\text{tot}} - p)} N_4)$, and $\Gamma_2(t) = \frac{1}{\sqrt{\Omega}}(\sqrt{\beta p} N_1 - \sqrt{\delta A} N_2)$, we can rewrite the above system in the following form:

$$\begin{aligned}
\frac{dp}{dt} &= -aAp + d(p_{\text{tot}} - p) + \sqrt{d}\Gamma_1(t) \\
\frac{dA}{dt} &= \beta p - \delta A - aAp + d(p_{\text{tot}} - p) + \Gamma_2(t) + \sqrt{d}\Gamma_1(t).
\end{aligned}$$

Since $a, d \gg \delta, \beta p$, this system displays two time scales. Denoting $\epsilon := \delta/d$ and defining $y := A - p$, the system can be further rewritten in standard singular perturbation form (3.6):

$$\begin{aligned}
\epsilon \frac{dp}{dt} &= -\delta Ap/K_{\text{d}} + \delta(p_{\text{tot}} - p) + \sqrt{\epsilon}\sqrt{\delta}\Gamma_1(t) \\
\frac{dy}{dt} &= \beta p - \delta(y + p) + \Gamma_2(t).
\end{aligned}$$

By setting $\epsilon = 0$ and assuming that $p_{\text{tot}}/K_{\text{d}}$ is sufficiently small, we obtain the reduced system describing the dynamics of A as

$$\frac{dA}{dt} = \beta \frac{p_{\text{tot}}}{A/K_{\text{d}} + 1} - \delta A + \frac{1}{\sqrt{\Omega}}(\sqrt{\beta p} N_1 - \sqrt{\delta A} N_2) =: f(A, N_1, N_2).$$

The equilibrium point for this system corresponding to the mean values $N_1 = 0$ and $N_2 = 0$ of the inputs is given by

$$A_e = \frac{1}{2}(\sqrt{K_{\text{d}}^2 + 4\beta p_{\text{tot}} K_{\text{d}}/\delta} - K_{\text{d}}).$$

The linearization of the system about this equilibrium point is given by

$$\left.\frac{\partial f}{\partial A}\right|_{A_e, N_1=0, N_2=0} = -\beta \frac{p_{\text{tot}}/K_{\text{d}}}{(A_e/K_{\text{d}} + 1)^2 + 1} - \delta =: -\delta,$$

$$b_1 = \left.\frac{\partial f}{\partial N_1}\right|_{A_e, N_1=0, N_2=0} = \frac{1}{\sqrt{\Omega}} \sqrt{\frac{\beta p_{\text{tot}}}{A_e/K_{\text{d}} + 1}}, \qquad b_2 = \left.\frac{\partial f}{\partial N_2}\right|_{A_e, N_1=0, N_2=0} = -\sqrt{\delta A_e}.$$

Hence, the frequency response to N_1 is given by

$$G^c_{AN1}(\omega) = \frac{b_1}{\sqrt{\omega^2 + \delta^2}}.$$

In order to make a fair comparison between this response and that of the open loop system, we need to make sure that the steady states of both systems are the same. In order to do so, we set

$$\beta_0 = \frac{\beta}{A_e/K_{\text{d}} + 1}.$$

This can be attained by properly adjusting the strength of the promoter and of the ribosome binding site.

As a consequence, $b_1 = \sqrt{\beta_0 p_{\text{tot}}/\Omega}$. Since also $\delta > \delta$, it is clear that $G^c_{AN1}(\omega) < G_{AN1}(\omega)$ for all ω. This result implies that the negative feedback attenuates the noise at all frequencies. The two frequency responses are plotted in Figure 6.2(a).

mRNA decay close to protein decay

In this case, we need to model the processes of transcription and translation separately. Denoting m_A the mRNA of A, the reactions describing the open loop system modify to

$$\text{R}_1\text{: m}_\text{A} \xrightarrow{\beta} \text{m}_\text{A} + \text{A}, \qquad \text{R}_2\text{: A} \xrightarrow{\delta} \emptyset, \qquad \text{R}_5\text{: p} \xrightarrow{\alpha} \text{m}_\text{A} + \text{p}, \qquad \text{R}_6\text{: m}_\text{A} \xrightarrow{\gamma} \emptyset,$$

while those describing the closed loop system modify to

$$\text{R}_1\text{: m}_\text{A} \xrightarrow{\beta} \text{m}_\text{A} + \text{A}, \qquad \text{R}_2\text{: A} \xrightarrow{\delta} \emptyset,$$

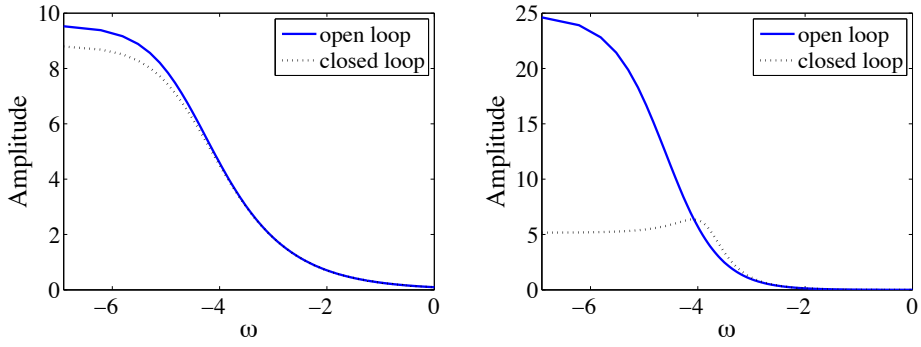


Figure 6.2: (a) Frequency response to noise $N_1(t)$ for both open loop and closed loop for the model in which mRNA is assumed at its quasi-steady state. The parameters are $p_{\text{tot}} = 10$, $K_{\text{d}} = 10$, $\beta = 0.001$, $\delta = 0.01$, $\Omega = 1$, and $\beta_0 = 0.00092$. (b) Frequency response to noise $N_6(t)$ for both open loop and closed loop for the model in which mRNA decay is close to protein decay. The parameters are $p_{\text{tot}} = 10$, $K_{\text{d}} = 10$, $\alpha = 0.001$, $\beta = 0.01$, $\gamma = 0.01$, $\delta = 0.01$, and $\alpha_0 = 0.0618$.

$$\text{R}_3\colon \text{A} + \text{p} \xrightarrow{a} \text{C}, \qquad \text{R}_4\colon \text{C} \xrightarrow{d} \text{A} + \text{p},$$

$$\text{R}_5\colon \text{p} \xrightarrow{\alpha} \text{m}_\text{A} + \text{p}, \qquad \text{R}_6\colon \text{m}_\text{A} \xrightarrow{\gamma} \emptyset, \qquad p_{\text{tot}} = p + C.$$

Employing the Langevin equation in terms of concentrations, and applying singular perturbation as performed before, we obtain the dynamics of the system as

$$\frac{dm_A}{dt} = f(A) - \gamma m_A + \frac{1}{\sqrt{\Omega}}\left(\sqrt{f(A)}N_5 - \sqrt{\gamma m_A}N_6\right)$$
$$\frac{dA}{dt} = \beta m_A - \delta A + \frac{1}{\sqrt{\Omega}}\left(\sqrt{\beta m_A}N_1 - \sqrt{\delta A}N_2\right),$$

in which N_5 and N_6 are the noise on the production reaction and decay reaction of mRNA, respectively. For the open loop system $f(A) = \alpha_0 p_{\text{tot}}$, while for the closed loop system

$$f(A) = \frac{\alpha p_{\text{tot}}}{A/K_{\text{d}} + 1}.$$

The steady state for the open loop system is given by

$$m_e^o = \frac{\alpha_0}{\gamma}, \qquad A_e^o = \frac{\alpha_0 \beta}{\gamma \delta}.$$

Considering N_6 to be the input of interest, the linearization of the system at this equilibrium is given by

$$A^o = \begin{pmatrix} -\gamma & 0 \\ \beta & -\delta \end{pmatrix}, \qquad B^o = \begin{pmatrix} \sqrt{\gamma m_e^o/\Omega} \\ 0 \end{pmatrix}.$$

Letting $K = \beta/(\delta K_{\rm d})$, the steady state for the closed loop system is given by

$$A_e^c = \frac{\beta m_e}{\delta}, \qquad m_e^c = \frac{1}{2}\left(-1/K + \sqrt{(1/K)^2 + 4\alpha p_{\rm tot}/(K\gamma)}\right).$$

The linearization of the closed loop system at this equilibrium point is given by

$$A^c = \begin{pmatrix} -\gamma & -G \\ \beta & -\delta \end{pmatrix}, \qquad B^c = \begin{pmatrix} \sqrt{\gamma m_e^c/\Omega} \\ 0 \end{pmatrix},$$

in which $G = \alpha p_{\rm tot}/(A_e^c/K_{\rm d} + 1)^2$ represents the contribution of the negative feedback. The larger the value of G the stronger the negative feedback.

In order to make a fair comparison between the two systems, we need to make the steady states be the same. In order to do this, we can set $\alpha_0 = \alpha/(A_e^c/K_{\rm d} + 1)$, which can be done by suitably changing the strengths of the promoter and ribosome binding sites.

The open loop and closed loop transfer functions are given by

$$G^o_{AN_6}(s) = \frac{\beta\sqrt{\gamma m_e/\Omega}}{(s+\gamma)(s+\delta)},$$

and by

$$G^c_{AN_6}(s) = \frac{\beta\sqrt{\gamma m_e/\Omega}}{s^2 + s(\gamma+\delta) + \gamma\delta + G},$$

respectively. By looking at these expressions, it is clear that the open loop transfer function has two real poles, while the closed loop transfer function can have complex conjugate poles when G is sufficiently large. As a consequence, noise N_6 can be amplified at sufficiently high frequencies. Figure 6.2(b) shows the corresponding frequency responses for both the open loop and the closed loop system.

It is clear that the presence of the negative feedback attenuates noise with respect to the open loop system at low frequency, but it amplifies it at higher frequency. This is a very well known effect known as the "water bed effect", according to which negative feedback decreases the effect of disturbances at low frequency, but it can amplify it at higher frequency. This effect is not found in first order models, as demonstrated by the derivations when mRNA is at the quasi-steady state. This illustrates the spectral shift of the frequency response to intrinsic noise toward the high frequency, as also experimentally demonstrated [6].

6.3 The Toggle Switch

The toggle switch is composed of two genes that mutually repress each other, as shown in the diagram of Figure 6.3 [31]. We start by describing a simple model

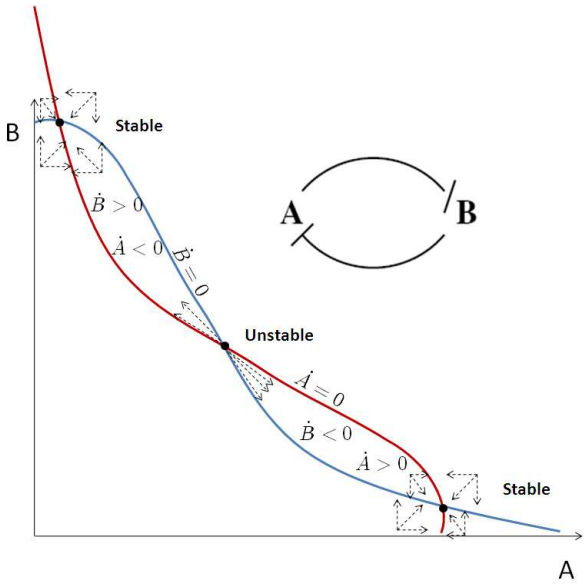


Figure 6.3: Nullclines for the toggle switch. By analyzing the direction of the vector field in the proximity of the equilibria, one can deduce their stability as described in Section 3.1.

with no inducers. By assuming that the mRNA dynamics are at the quasi-steady state, we obtain a two dimensional differential equation model given by

$$\frac{dA}{dt} = \frac{\beta}{1+(B/K)^n} - \delta A, \qquad \frac{dB}{dt} = \frac{\beta}{1+(A/K)^n} - \delta B,$$

in which we have assumed for simplicity that the parameters of the repression functions are the same for A and B.

The number and stability of equilibria can be analyzed by performing nullcline analysis since the system is two-dimensional. Specifically, by setting $dA/dt = 0$ and $dB/dt = 0$, we obtain the nullclines shown in Figure 6.3. In the case in which the parameters are the same for both A and B, the nullclines intersect at three points, which determine the steady states of this system.

The nullclines partition the plane into six regions. By determining the sign of dA/dt and dB/dt in each of these six regions, one determines the direction in which the vector field is pointing in each of these regions. From these directions, one immediately deduces that the steady state for which $A = B$ is unstable while the other two are stable. This is thus a bistable system.

The system converges to one steady state or the other depending on the initial condition. If the initial condition is in the region of attraction of one steady state, it converges to that steady state. The 45 degree line divides the plane into the two regions of attraction of the stable steady states. Once the system has converged to one of the two steady states, it cannot switch to the other unless an external

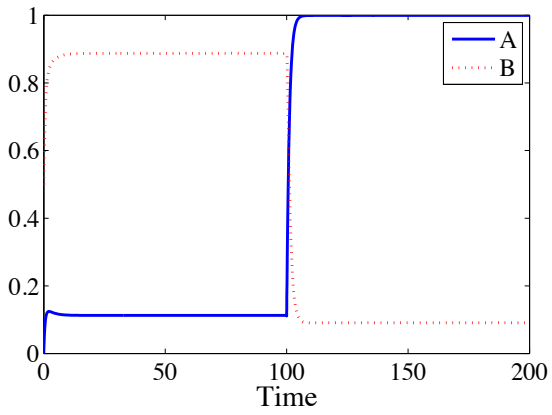


Figure 6.4: time traces for $A(t)$ and $B(t)$ when inducer concentrations u_1 and u_2 are changed. In the simulation, we have $n = 2$, $K_{d,1} = K_{d,2} = 1$, $K^2 = 0.1$, $\beta = 1$, and $\delta = 1$. The inducers are such that $u_1 = 10$ for $t < 100$ and $u_1 = 0$ for $t \geq 100$, while $u_2 = 0$ for $t < 100$ and $u_2 = 10$ for $t \geq 100$.

stimulation is applied that moves the initial condition to the region of attraction of the other steady state.

In the toggle switch by [31], external stimulations were added in form of negative inducers for A and B. Let u_1 be the negative inducer for A and u_2 be the negative inducer for B. Then, as we have seen in Section 2.3, the expressions of the Hill functions need to be modified to replace A by $A(1/(1+u_1/K_{d,1}))$ and B by $B(1/(1+u_2/K_{d,2}))$, in which $K_{d,1}$ and $K_{d,2}$ are the dissociation constants of u_1 with A and of u_2 with B, respectively. We show in Figure 6.4 time traces for $A(t)$ and $B(t)$ when the inducer concentrations are changed. Specifically, initially u_1 is high until time 100 while u_2 is low until this time. As a consequence, A does not repress B while B represses A. Accordingly, the concentration of A stays low until time 100 and the concentration of B stays high. After time 100, u_2 is high and u_1 is low. As a consequence B does not repress A while A represses B. In this situation, A switches to its high value and B switches to its low value.

6.4 The Repressilator

Elowitz and Leibler [27] constructed the first operational oscillatory genetic circuit consisting of three repressors arranged in ring fashion, and coined it the "repressilator" (Figure 6.1d). The repressilator exhibits sinusoidal, limit cycle oscillations in periods of hours, slower than the cell-division life cycle. Therefore, the state of the oscillator is transmitted between generations from mother to daughter cells.

The dynamical model of the repressilator can be obtained by composing three

transcriptional modules in a loop fashion. The dynamics can be written as

$$\frac{dm_A}{dt} = -\delta m_A + f_1(C) \qquad \frac{dm_B}{dt} = -\delta m_B + f_2(A) \qquad \frac{dm_C}{dt} = -\delta m_C + f_3(B)$$
$$\frac{dA}{dt} = m_A - \delta A \qquad \frac{dB}{dt} = m_B - \delta B \qquad \frac{dC}{dt} = m_C - \delta C,$$

where we take

$$f_1(p) = f_2(p) = f_3(p) = \frac{\alpha^2}{1+p^n}.$$

This structure belongs to the class of cyclic feedback systems that we have studied in Section 3.4. In particular, the Mallet-Paret and Smith theorem and Hastings theorem (see Section 3.4 for the details) can be applied to infer that if the system has a unique equilibrium point and this is unstable, then it admits a periodic solution. Therefore, we first determine the number of equilibria and their stability.

The equilibria of the system can be found by setting the time derivatives to zero. We thus obtain that

$$A = \frac{f_1(C)}{\delta^2}, \qquad B = \frac{f_2(A)}{\delta^2}, \qquad C = \frac{f_3(B)}{\delta^2},$$

which combined together yield to

$$A = \frac{1}{\delta^2} f_1\left(\frac{1}{\delta^2} f_3\left(\frac{1}{\delta^2} f_2(A)\right)\right) =: g(A).$$

The solution to this equation determines the set of steady states of the system. The number of steady states is given by the number of crossings of the two functions $h_1(A) = g(A)$ and $h_2(A) = A$. Since h_2 is strictly monotonically increasing, we obtain a unique steady state if h_1 is monotonically decreasing. This is the case when $g'(A) = \frac{dg(A)}{dA} < 0$. Otherwise, there could be multiple steady states. Since we have that

$$\text{sign}(g'(A)) = \Pi_{i=1}^3 \text{sign}(f_i'(P)),$$

then if $\Pi_{i=1}^3 \text{sign}(f_i'(P)) < 0$ the system has a unique steady state. We call the product $\Pi_{i=1}^3 \text{sign}(f_i'(P))$ the *loop gain*.

Thus, any cyclic feedback system with negative loop gain will have a unique steady state. It can be shown that a cyclic feedback system with positive loop gain belongs to the class of monotone systems and hence cannot have periodic orbits [62]. In the present case, system (6.4) is such that $f_i' < 0$, so that the loop gain is negative and there is a unique steady state. We next study the stability of this steady state by studying the linearization of the system.

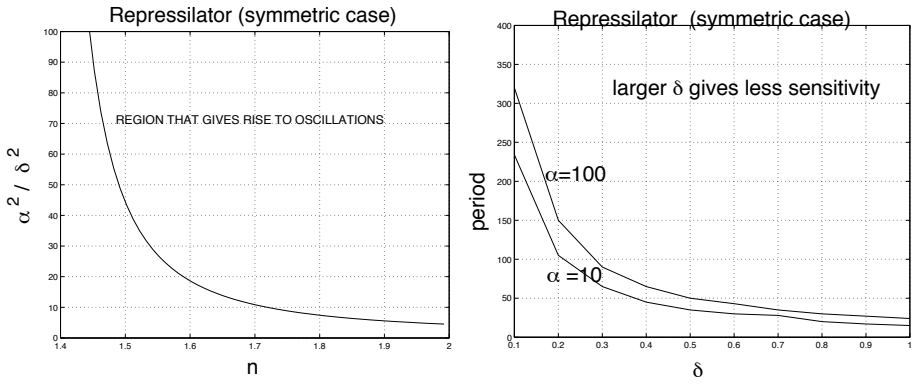


Figure 6.5: (a) Space of parameters that give rise to oscillations for the repressilator in equation (6.4). (b) Period as a function of δ and α.

Letting P denote the steady state value of the protein concentrations for A, B, and C, the linearization of the system is given by

$$J = \begin{pmatrix} -\delta & 0 & 0 & 0 & 0 & f_1'(P) \\ 1 & -\delta & 0 & 0 & 0 & 0 \\ 0 & f_2'(P) & -\delta & 0 & 0 & 0 \\ 0 & 0 & 1 & -\delta & 0 & 0 \\ 0 & 0 & 0 & f_3'(P) & -\delta & 0 \\ 0 & 0 & 0 & 0 & 1 & -\delta \end{pmatrix},$$

whose characteristic polynomial is given by

$$\det(\lambda I - J) = (\lambda + \delta)^6 - \Pi_{i=1}^3 f_i'(P). \tag{6.1}$$

In the case in which $f_i(P) = \alpha^2/(1+p^n)$ for $i \in \{1,2,3\}$, this characteristic polynomial has a root with positive real part if the ratio α/δ satisfies the relation

$$\alpha^2/\delta^2 > \sqrt[n]{\frac{4/3}{n-4/3}}\left(1+\frac{4/3}{n-4/3}\right).$$

For the proof of this statement, the reader is referred to [21]. This relationship is plotted in Figure 6.5 (b).

When n increases, the existence of an unstable equilibrium point is guaranteed for larger ranges of the other parameter values. Of course, this "behavioral" robustness does not guarantee that other important features of the oscillator, such as the period are not changed when parameters vary. Numerical studies indicate that the period T approximatively follows $T \propto 1/\delta$, and varies little with respect to α (Figure 6.5b). From the figure, we see that as the value of δ increases, the sensitivity of

the period to the variation of δ itself decreases. However, increasing δ would necessitate the increase of the cooperativity n, therefore indicating a possible tradeoff that should be taken into account in the design process in order to balance the system complexity and robustness of the oscillations. From a practical point of view, n can be changed by selecting repressors that bind cooperatively to the promoter. In practice, it is usually hard to obtain values of n greater than two.

A similar result for the existence of a periodic solution can be obtained for the non-symmetric case in which the input functions of the three transcriptional modules are modified to

$$f_1(p) = \frac{\alpha_3^2}{1+p^n}, \qquad f_2(p) \quad = \frac{\alpha^2 p^n}{1+p^n}, \qquad f_3(p) = \frac{\alpha^2 p^n}{1+p^n}.$$

That is, two interactions are activations and one only is a repression. Since the loop gain is still negative, there is only one equilibrium point only. We can thus obtain the condition for oscillations again by establishing conditions on the parameters that guarantee that at least one root of the characteristic polynomial (6.1) has positive real part, that is,

$$(0.86)^2 n \sqrt[3]{\frac{p_3^n}{(1+p_3^n)(1+p_2^n)(1+p_1^n)}} > 1. \tag{6.2}$$

We rewrite p_1 and p_3 as functions of p_2 by using two of the equilibrium relations:

$$p_1 = \sqrt[n]{\frac{p_2}{\alpha^2/\delta^2 - p_2}}, \qquad p_3 = \frac{\alpha^2/\delta^2 p_2^n}{1+p_2^n}.$$

Using these expressions in (6.2), we can find all possible values of p_2 that satisfy (6.2) for a fixed pair $(\alpha^2/\delta^2, n)$. These values of p_2 correspond to the possible values of α_3^2/δ^2 by means of the third equilibrium condition

$$\alpha_3^2/\delta^2 = p_1(1+p_3^n).$$

For each pair $(\alpha^2/\delta^2, n)$, we finally obtain all possible values of α_3^2/δ^2 that satisfy the equilibrium conditions and inequality (6.2). These conditions are reported in Figure 6.6 (see [21] for the detailed derivations).

One can conclude that it is possible to "over design" the circuit to be in the region of parameter space that gives rise to oscillations. In practice, values of n between one and two can be obtained by employing repressors that have coopearivity higher than or equal to two. There are plenty of such repressors, including those originally used in the repressilator design [27]. However, values of n greater than two may be hard to reach in practice. It is also possible to show that increasing the number of elements in the loop, the value of n sufficient for oscillatory behavior decreases (see Exercises).

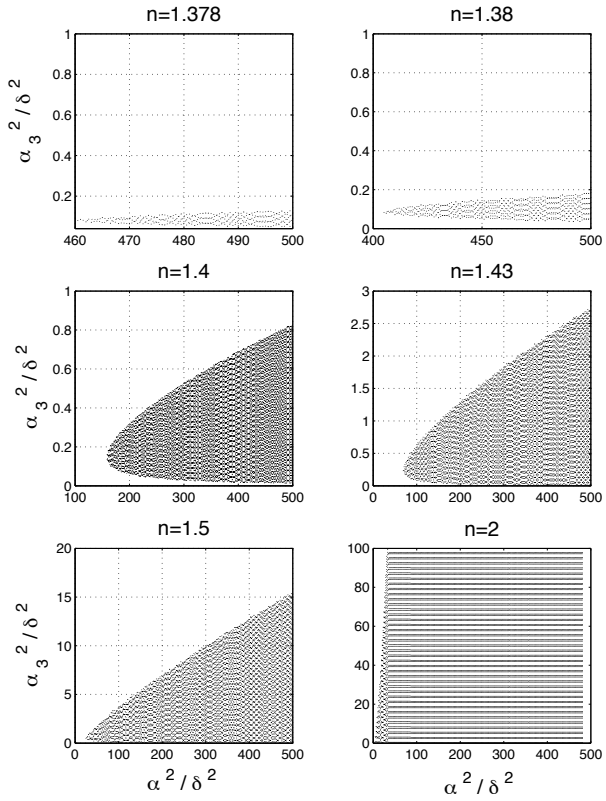
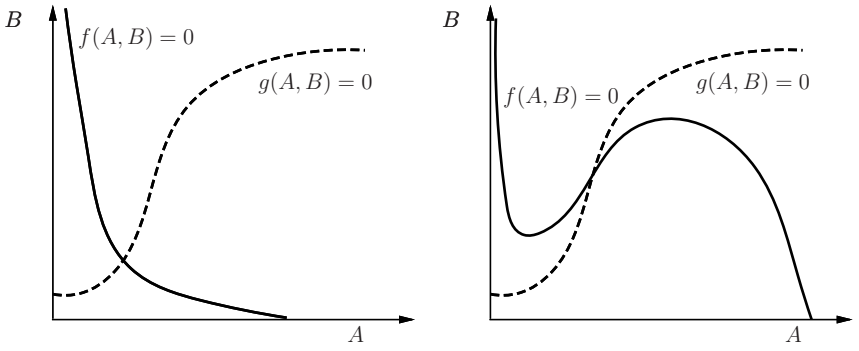


Figure 6.6: Space of parameters that give rise to oscillations for the repressilator (non-symmetric case). As the value of n is increased, the ranges of the other parameters for which sustained oscillations exist become larger.

6.5 Activator-repressor clock

Consider the activator-repressor clock diagram shown in Figure 6.1(c). The transcriptional module A has an input function that takes two inputs: an activator A and a repressor B. The transcriptional module B has an input function that takes only an activator A as its input. Let m_A and m_B represent the concentration of mRNA of the activator and of the repressor, respectively. Let A and B denote the protein concentration of the activator and of the repressor, respectively. Then, we consider the following four-dimensional model describing the rate of change of the species concentrations:

$$\frac{dm_A}{dt} = -\delta_1 m_A + F_1(A,B), \qquad \frac{dm_B}{dt} = -\delta_2 r_B + F_2(A),$$
$$\frac{dA}{dt} = -\delta_A A + \beta_1 m_A, \qquad \frac{dB}{dt} = -\delta_B B + \beta_2 m_B,$$



(a) $n = 1$ (b) $n = 2$

Figure 6.7: Nullclines for the two-dimensional system of equation (6.5). (a) shows the only possible configuration of the nullclines when $n = 1$. (b) shows a possible configuration of the nullclines when $n = 2$. In this configuration, there is a unique equilibrium, which can be unstable.

in which the functions F_1 and F_2 are Hill functions and given by

$$F_1(A,B) = \frac{K_1 A^n + K_{A0}}{1 + (A/k_1)^n + (B/k_2)^m}, \quad F_2(A) = \frac{K_2 A^n + K_{B0}}{1 + (A/k_1)^n}.$$

The Hill function F_1 can be obtained through a combinatorial promoter, where there are sites both for an activator and for a repressor (see Section 2.3).

Two-dimensional analysis

We first assume the mRNA dynamics to be at the quasi-steady state so that we can perform two dimensional analysis and invoke the Poincarè-Bendixson theorem. Then, we analyze the four dimensional system and perform a bifurcation study.

We let $f_1(A,B) := (\beta_1/\delta_1)F_1(A,B)$ and $f_2(A) := (\beta_2/\delta_2)F_2(A)$. For simplicity, we also denote $f(A,B) := -\delta_A A + f_1(A,B)$ and $g(A,B) := -\delta_B B + f_2(A)$ so that the two-dimensional system is given by

$$\frac{dA}{dt} = f(A,B), \qquad \frac{dB}{dt} = g(A,B).$$

For simplicity, we assume $m = 1$ and $k_i = 1$ for all i.

We first study whether the system admits a periodic solution for $n = 1$. To do so, we analyze the nullclines to determine the number and location of steady states. Denote $K_1 = K_1(\beta_1/\delta_1)$, $K_2 = K_2(\beta_2/\delta_2)$, $K_{A0} = K_{A0}(\beta_1/\delta_1)$, and $K_{B0} = K_{B0}(\beta_1/\delta_1)$. Then, $g(A,B) = 0$ leads to

$$B = \frac{K_2 A + K_{B0}}{(1+A)\delta_A},$$

which is an increasing function of A. Setting $f(A,B)=0$, we obtain that

$$B=\frac{K_1A+K_{A0}-\delta_A A(1+A)}{\delta_A A},$$

which is a monotonically decreasing function of A. These nullclines are displayed in Figure 6.7(a).

We see that we have one equilibrium only. To determine the stability of such an equilibrium, we calculate the linearization of the system at such an equilibrium. This is given by

$$J=\begin{pmatrix} \frac{\partial f}{\partial A} & \frac{\partial f}{\partial B} \\ \frac{\partial g}{\partial A} & \frac{\partial g}{\partial B} \end{pmatrix}$$

In order for the equilibrium to be unstable and not a saddle, it is necessary and sufficient that $\mathrm{tr}(J)>0$ and $\det(J)>0$.

Graphical inspection of the nullclines at the equilibrium (see 6.7(a)), shows that

$$\left.\frac{dB}{dA}\right|_{f(A,B)=0}<0.$$

By the implicit function theorem (Section 3.6), we further have that

$$\left.\frac{dB}{dA}\right|_{f(A,B)=0}=-\frac{\partial f/\partial A}{\partial f/\partial B},$$

so that $\partial f/\partial A<0$ because $\partial f/\partial B<0$. As a consequence, we have that $\mathrm{tr}(J)<0$ and hence the equilibrium point is either stable or a saddle.

To determine the sign of det(J), we further inspect the nullclines and find that

$$\left.\frac{dB}{dA}\right|_{g(A,B)=0}>\left.\frac{dB}{dA}\right|_{f(A,B)=0}.$$

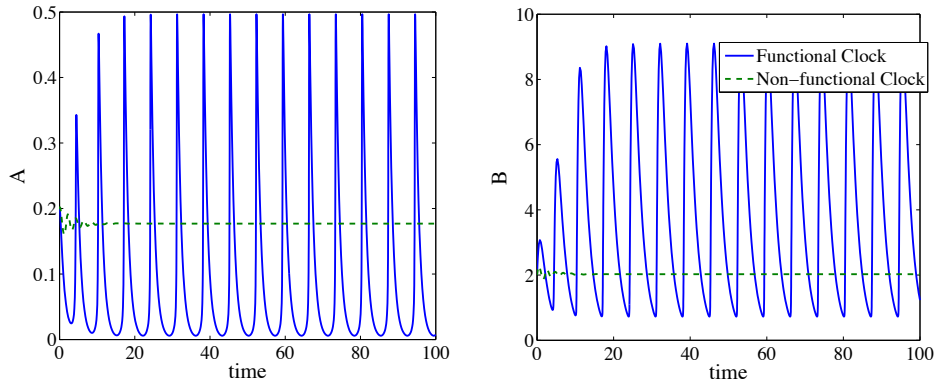
Again using the implicit function theorem we have that

$$\left.\frac{dB}{dA}\right|_{g(A,B)=0}=-\frac{\partial g/\partial A}{\partial g/\partial B},$$

so that $\det(J)>0$. Hence, the ω-limit set (Section 3.4) of any point in the plane is not necessarily a periodic orbit. Hence, to guarantee that any initial condition converges to a periodic orbit, we need to require that $n>1$.

We now study the case $n=2$. In this case, the nullcline $f(A,B)=0$ changes and can have the shape shown in Figure 6.7 (b). In the case in which, as in the figure, there is an equilibrium point only and the nullclines intersect both with positive slope (equivalent to $\det(J)>0$), the equilibrium is unstable and not a saddle if $\mathrm{tr}(J)>0$, which is satisfied if

$$\frac{\delta_B}{\partial f_1/\partial A-\delta_A}<1.$$



(a) (b)

Figure 6.8: *Effect of the trace of the Jacobian on the stability of the equilibrium*. The above plots illustrate the trajectories of system (6.5) for both Functional ($\mathrm{tr}(J) > 0$) and a Non-Functional ($\mathrm{tr}(J) < 0$) Clocks. The parameters in the simulation are $\delta_1 = \delta_2 = 1$, $K_1 = K_2 = 100$, $K_{A0} = .04$, $K_{B0} = .004$, $\delta_A = 1$, $\beta_1 = \beta_2 = 1$, and $k_1 = k_2 = 1$. In the Functional Clock, $\delta_B = 0.5$ whereas in the Non-Functional Clock, $\delta_B = 1.5$. Parameters K_1 and K_2 were chosen to give about 500-2000 copies of protein per cell for activated promoters. Parameters K_{A0} and K_{B0} were chosen to give about 1-10 copies per cell for non-activated promoters.

This condition reveals the crucial design requirement for the functioning of the clock. Specifically the repressor B time scale must be sufficiently slower than the activator A time scale. This point is illustrated in the simulations of Figure 6.8, in which we see that if δ_B is too large, the trace becomes negative and oscillations disappear.

Four-dimensional analysis

In order to specifically study time scale separation between activator and repressor as a crucial design requirement for the clock, we perform a time scale analysis employing bifurcation the tools described in Section 3.5. To this end, we consider the following four-dimensional model describing the rate of change of the species concentrations:

$$\frac{dm_A}{dt} = -\delta_1/\epsilon\, m_A + F_1(A,B), \qquad \frac{dm_B}{dt} = -\delta_2/\epsilon\, m_B + F_2(A),$$
$$\frac{dA}{dt} = \nu(-\delta_A A + \beta_1/\epsilon\, m_A), \qquad \frac{dB}{dt} = -\delta_B B + \beta_2/\epsilon\, m_B.$$

This system is the same as system (6.5) where we have explicitly introduced two parameters, ν and ϵ, which model time scale differences as follows. The parameter ν regulates the difference of time scale between the repressor and the activator

dynamics while ϵ regulates the difference of time scale between the mRNA and the protein dynamics. The parameter ϵ determines how close model (6.5) is to the two-dimensional model (6.5), in which the mRNA dynamics are considered at the quasi-steady state. Thus, ϵ is a singular perturbation parameter (equations (6.5) can be taken to standard singular perturbation form by considering the change of variables $\overline{m}_A = m_A/\epsilon$ and $\overline{m}_B = m_B/\epsilon$). The details on singular perturbation can be found in Section 3.6.

The values of ϵ and of ν do not affect the number of equilibria of the system. We then perform bifurcation analysis with ϵ and ν the two bifurcation parameters. The bifurcation analysis results are summarized by Figure 6.9. The reader is referred to [20] for the details of the numerical analysis. In terms of the ϵ and ν parameters, it is thus possible to "over design" the system: if the activator dynamics are sufficiently sped up with respect to the repressor dynamics, the system undergoes a Hopf bifurcation (Hopf bifurcation was introduced in Section 3.4) and stable oscillations will arise.

From a fabrication point of view, the activator dynamics can be sped up by adding suitable degradation tags to the activator protein. Similarly, the repressor dynamics can be slowed down by adding repressor DNA binding sites (see Chapter 7 and the effects of retroactivity on dynamic behavior).

6.6 An Incoherent Feedforward Loop (IFFL)

Several genetic implementations of incoherent feedforward loops are possible [3]. Here, we describe an implementation proposed for making the steady state levels of protein expression adapt to DNA plasmid copy number [13]. In this implementation, the input u is the amount of DNA plasmid coding for both the intermediate regulator LacI (L) with concentration L and the output RFP (R) with concentration R. The intermediate regulator LacI represses through transcriptional repression the expression of the output protein RFP (Figure 6.10). The expectation is that the steady state value of the RFP expression is independent of the concentration u of the plasmid. That is, the concentration of RFP should adapt to the copy number of its own plasmid.

In order to analyze whether the adaptation property holds, we write the differential equation model describing the system, assuming the mRNA dynamics are at the quasi-steady state. This model is given by

$$\frac{dL}{dt} = k_0 u - \delta L, \qquad \frac{dR}{dt} = \frac{k_1 u}{1 + (L/K_{\mathrm{d}})} - \delta R, \tag{6.3}$$

in which k_0 is the constitutive rate at which LacI is expressed and K_{d} is the dissociation constant of LacI from the operator sites on the lac promoter. This implementation has been called the sniffer in Section 3.2. The steady state of the system is

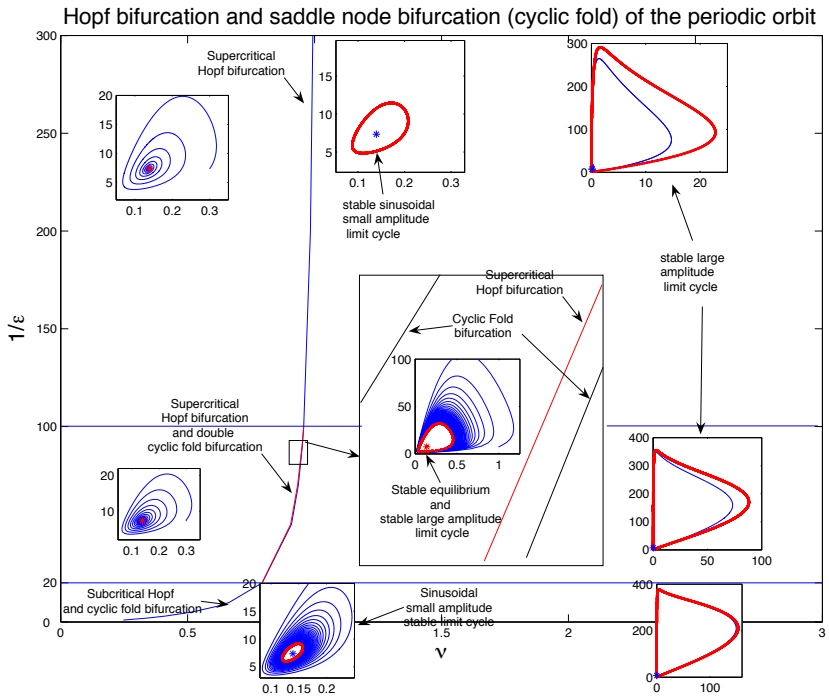


Figure 6.9: Design chart for the relaxation oscillator. We obtain sustained oscillations past the Hopf bifurcation point, for values of ν sufficiently large independently of the difference of time scales between the protein and the mRNA dynamics. We also notice that there are values of ν for which a stable equilibrium point and a stable orbit coexist and values of ν for which two stable orbits coexist. The interval of ν values for which two stable orbits coexist is too small to be able to numerically set ν in such an interval. Thus, this interval is not practically relevant. The values of ν for which a stable equilibrium and a stable periodic orbit coexist is instead relevant. This situation corresponds to the *hard excitation* condition [58] and occurs for realistic values of the separation of time-scales between protein and m-RNA dynamics. Therefore, this simple oscillator motif described by a four-dimensional model can capture the features that lead to the long term suppression of the rhythm by external inputs.

obtained by setting the time derivatives to zero and gives

$$L = \frac{k_0}{\delta} u, \qquad R = \frac{k_1 u}{\delta + k_0 u / K_\mathrm{d}}.$$

From this expression, one can easily note that as K_d decreases, the denominator of the right-side expression tends to $k_0 u / K_\mathrm{d}$ resulting into the steady state value $R = k_1 K_\mathrm{d} / k_0$, which does not depend on the input u. Hence, in this case, adaptation would be reached. This is the case if the affinity of LacI to its operator sites is extremely high, resulting also in a strong repression and hence a lower value of R. In practice, however, the value of K_d is non-zero, hence the adaptation is not

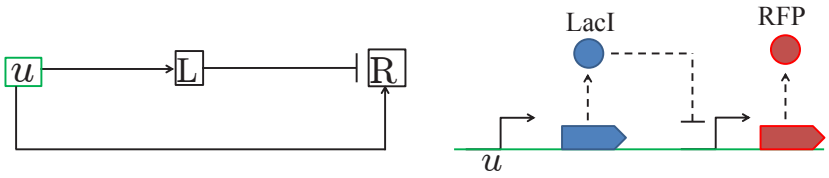


Figure 6.10: (Left) The incoherent feedforward motif. (Right) A possible implementation of the incoherent feedforward motif. Here, LacI (L) is under the control of a constitutive promoter in amounts u, while RFP (R) is under the control of the lac promoter, also in amounts u. Hence RFP is also activated by u as the RFP gene is found in amounts u just like the LacI gene.

perfect. We show in Figure 6.11 the behavior of the steady state of R as a function of the input u for different values of K_d. Ideally, for perfect adaptation, this should be a horizontal line.

In this study, we have modeled protein L as binding with its promoter with no cooperativity. If L is LacI, the cooperativity of binding is $n = 4$. We leave as an exercise to show that the adaptation behavior persist in this case (see Exercises).

For engineering a system with prescribed behavior, one has to be able to change the physical features so as to change the values of the parameters of the model. This is often possible. For example, the binding affinity ($1/K_d$ in the Hill function model) of a transcription factor to its site on the promoter can be affected by single or multiple base pairs substitutions. The protein decay rate can be increased by adding degradation tags at the end of the gene expressing protein Y. Promoters that can accept multiple input transcription factors (combinatorial promoters) to implement regulation functions that take multiple inputs can be realized by combining the operator sites of several simple promoters [19].

Exercises

6.1 Consider the toggle switch:

$$\dot{A} = \frac{\beta}{1+(B/K_1)^n} - \alpha_1 A, \qquad \dot{B} = \frac{\gamma}{1+(A/K_2)^m} - \alpha_2 B.$$

Here, we are going to explore the parameter space that makes the system work as a toggle. To do so, answer the following questions:

(a) Consider $m = n = 1$. Determine the number and stability of the equilibria.

(b) Consider $m = 1$ and $n > 1$ and determine the number and stability of the equilibria (as other parameters change).

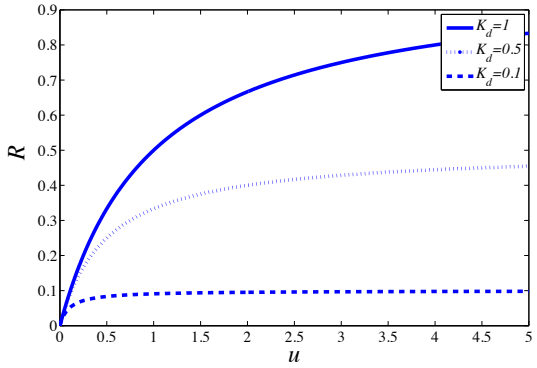


Figure 6.11: Behavior of the steady state value of y as a function of the input u.

(c) Consider $m = n = 2$. Determine parameter conditions on $\beta, \gamma, \alpha_1, \alpha_2$ for which the system is bistable, i.e., there are two stable steady states.

6.2 Consider the "generalized" model of the repressilator in which we have m repressors (with m an odd number) in the ring. Explore via simulation the fact that when m is increased, the system oscillates for smaller values of the Hill coefficient n.

6.3 Consider the oscillator design of Stricker et al. [90]. Build a four dimensional model including mRNA concentration and protein concentration. Then reduce this fourth order model to a second order model using the QSS approximation for the mRNA dynamics. Then, investigate the following points:

(a) Use the Poincaré-Bendixson theorem to determine under what conditions the system in 2D admits a periodic orbit.

(b) Simulate the four dimensional system and the two dimensional system for parameter values that give oscillations and study how close the trajectories of the 2D approximation are to those of the 4D system.

(c) Determine whether the four dimensional system has a Hopf bifurcation (either analytically or numerically).

6.4 Consider the feedforward circuit shown in Figure 6.11. Assume now to model the fact that the cooperativity of binding of LacI to its promoter is 4. The model then modifies to

$$\frac{dL}{dt} = k_0 u - \delta L, \qquad \frac{dR}{dt} = \frac{k_1 u}{1 + (L/K_\mathrm{d})^4} - \delta R. \tag{6.4}$$

Show that the adaptation property still holds under suitable parameter conditions.