Biomolecular Feedback Systems

Domitilla Del Vecchio
MIT

Richard M. Murray
Caltech

Classroom Copy v0.6d, July 20, 2013

Contents

Preface

This text is intended for researchers interested in the application of feedback and control to biomolecular systems. The material has been designed so that it can be used in parallel with the textbook *Feedback Systems* [1] as part of a course on biomolecular feedback and control systems, or as a standalone reference for readers who have had a basic course in feedback and control theory. The full text for this book, along with additional supplemental material, is available on a companion web site:

`http://www.cds.caltech.edu/~murray/BFS`

The material in this book is intended to be useful to three overlapping audiences: graduate students in biology and bioengineering interested in understanding the role of feedback in natural and engineered biomolecular systems; advanced undergraduates and graduate students in engineering disciplines who are interested the use of feedback in biological circuit design; and established researchers in the biological sciences who want to explore the potential application of principles and tools from control theory to biomolecular systems. We have written the text assuming some familiarity with basic concepts in feedback and control, but have tried to provide insights and specific results as needed, so that the material can be learned in parallel. We also assume some familiarity with cell biology, at the level of a first course for non-majors. The individual chapters in the text indicate the pre-requisites in more detail, most of which are covered either in AM08 or in the supplemental information available from the companion web site.

Domitilla Del Vecchio
Cambridge, Massachusetts

Richard M. Murray
Pasadena, California

Notation

This is an internal chapter that is intended for use by the authors in fixing the notation that is used throughout the text. In the first pass of the book we are anticipating several conflicts in notation and the notes here may be useful to early users of the text.

Protein dynamics

For a gene 'genX', we write *genX* for the gene, $\mathrm{m}_{\mathrm{genX}}$ for the mRNA and GenX for the protein when they appear in text or chemical formulas. Superscripts are used for covalent modifications, e.g., X^{p} for phosphorylation. We also use superscripts to differentiate between isomers, so $\mathrm{m}^{*}_{\mathrm{genX}}$ might be used to refer to mature RNA or $\mathrm{GenX}^{\mathrm{f}}$ to refer to the folded versions of a protein, if required. Mathematical formulas use the italic version of the variable name, but roman font for the gene or isomeric state. The concentration of mRNA is written in text or formulas as m_{genX} (m^{*}_{genX} for mature) and the concentration of protein as p_{genX} ($p^{\mathrm{f}}_{\mathrm{genX}}$ for folded). The same naming conventions are used for common gene/protein combinations: the mRNA concentration of *tetR* is $\mathrm{m}_{\mathrm{tetR}}$, the concentration of the associated protein is p_{tetR} and parameters are α_{tetR}, γ_{tetR}, etc.

For generic genes and proteins, use X to refer to a protein, m_{x} to refer to the mRNA associated with that protein and *x* to refer to the gene that encodes X. The concentration of X can be written either as X, p_{x} or [X], with that order of preference. The concentration of m_{x} can be written either as m_{x} (preferred) or $[\mathrm{m}_{\mathrm{x}}]$. Parameters that are specific to gene *p* are written with a subscripted p: α_{p}, γ_{p}, etc. Note that although the protein is capitalized, the subscripts are lower case (so indexed by the gene, not the protein) and also in roman font (since they are not a variable).

Transcription and translation. The dynamics of protein production are given by

$$\frac{dm_P}{dt} = \alpha_P \underbrace{- \mu m_P - \bar{\delta}_P m_P,}_{-\delta_P m_P} \qquad \frac{dP}{dt} = \kappa_P m_P \underbrace{- \mu P - \bar{\gamma}_P P,}_{-\gamma_P P}$$

where $\alpha_{P,0}$ is the (basal) rate of production, δ_P parameterizes the rate of degradation of the mRNA m_{p}, β_P is the kinetic rate of protein production, μ is the growth rate that leads to dilution of concentrations and γ_P parameterizes the rate of degradation

of the protein P. Since dilution and degradation enter in a similar fashion, we use $\delta = \bar{\delta} + \mu$ and $\gamma = \bar{\gamma} + \mu$ to represent the aggregate degradation and dilution rate. If we are looking at a single gene/protein, the various subscripts can be dropped.

When we ignore the mRNA concentration, we write the simplified protein dynamics as

$$\frac{dP}{dt} = \beta_P - \gamma_P P.$$

Assuming that the mRNA dynamics are fast compared to protein production, then the constant β_P is given by

$$\beta_P = \kappa_P \frac{\alpha_P}{\delta_P}.$$

In general, if this does not create confusion, we remove the subscripts "P" from the parameters.

Hill functions. For regulated production of proteins using Hill functions, we modify the constitutive rate of production to be $F(Q)$, in which Q is a transcription factor, instead of α_p or β_p as appropriate. The Hill function is written in the forms

$$F(Q) = \frac{\alpha}{1 + (Q/K)^n}, \qquad F(Q) = \frac{\alpha (Q/K)^n}{1 + (Q/K)^n} + \alpha_0.$$

The notation for F mirrors that of transfer functions in AM08: $F_{\mathrm{p,q}}$ represents the input/output relationship between input Q and output P (rate). If the target gene is not particularly relevant, the subscript can represent just the transcription factor:

$$F_\mathrm{lac}(Q) = \frac{\alpha_\mathrm{lac}}{1 + (Q/K_\mathrm{laq})^{n_\mathrm{lac}}}.$$

The subscripts can be dropped completely if there is only one Hill function in use.

Concentrations. For a species A, A is its concentration, that is, A:=[A]. n_A is the number of A molecules and m_A is the mRNA.

For complexes ES (complex of E and S), we denote $C = [ES]$ and write differential equations with C only or [ES], that is, $\frac{dC}{dt}$ or $\frac{d[ES]}{dt}$.

For names of proteins, such as TetR, we write T :=[TetR] and everything follows the rules of the species A.

Vector fields. $\dot{x} = f(x)$ or $\dot{x} = f(x, u, \theta)$: all are lower case. Upper case F is reserved for Hill functions.

Some common symbols:

Symbol	LaTeX	Comment
X_tot	`X_\tot`	Total concentration of a species
K_d	`\Kd`	Dissociation constant
K_m	`\Km`	Michaelis-Menten constant

Chemical reactions

We write the symbol for a chemical species A using roman type. The number of molecules of a species A is written as n_a. The concentration of the species is occasionally written as [A], but we more often use the notation A, as in the case of proteins, or x_a. For a reaction $\mathrm{A} + \mathrm{B} \longleftrightarrow \mathrm{C}$, we use the notation

$$\texttt{R1}: \quad \mathrm{A} + \mathrm{B} \underset{d_1}{\overset{a_1}{\rightleftharpoons}} \mathrm{C} \qquad \frac{dC}{dt} = a_1 AB - d_1 C$$

This notation is primarily intended for situations where we have multiple reactions and need to distinguish between many different constants. Enzymatic reactions have the form

$$\texttt{R2}: \quad \mathrm{S} + \mathrm{E} \underset{d_2}{\overset{a_2}{\rightleftharpoons}} \mathrm{C} \xrightarrow{k} \mathrm{P} + \mathrm{E}$$

For a small number of reactions, the reaction number can be dropped.

It will often be the case that two species A and B will form a molecular bond, in which case we write the resulting species as AB. If we need to distinguish between covalent bonds and hydrogen bonds, we write the latter as A:B. Finally, in some situations we will have labeled section of DNA that are connected together, which we write as A–B, where here A represents the first portion of the DNA strand and B represents the second portion. When describing (single) strands of DNA, we write A′ to represent the Watson-Crick complement of the strand A. Thus A–B:B′–A′ would represent a double stranded length of DNA with domains A and B.

The choice of representing covalent molecules using the conventional chemical notation AB can lead to some confusion when writing the reaction dynamics using A and B to represent the concentrations of those species. Namely, the symbol AB could represent either the concentration of A times the concentration of B or the concentration of AB. To remove this ambiguity, when using this notation we write [A][B] as $A \cdot B$.

When working with a system of chemical reactions, we write S_i, $i = 1, \ldots, n$ for the species and R_j, $j = 1, \ldots, m$ for the reactions. We write n_i to refer to the molecular count for species i and $x_i = [\mathrm{S}_i]$ to refer to the concentration of the species. The individual equations for a given species are written

$$\frac{dx_i}{dt} = \sum_{j=1}^{m} k_{i,jk}\, x_j x_k.$$

The collection of reactions are written as

$$\frac{dx}{dt} = N v(x, \theta), \qquad \frac{dx_i}{d =} N_{ij} v_j(x, \theta),$$

where x_i is the concentration of species S_i, $N \in \mathbb{R}^{n \times m}$ is the stochiometry matrix, v_j is the reaction flux vector for reaction j, and θ is the collection of parameters that

the define the reaction rates. Occasionally it will be useful to write the fluxes as polynomials, in which case we use the notation

$$v_j(x,\theta) = \sum_k E_{jk} \prod_l x_l^{\epsilon_l^{jk}}$$

where E_{jk} is the rate constant for the kth term of the jth reaction and ϵ_l^{jk} is the stochiometry coefficient for the species x_l.

Generally speaking, coefficients for propensity functions and reaction rate constants are written using lower case (c_ξ, k_i, etc). Two exceptions are the dissociation constant, which we write as K_d, and the Michaelis-Menten constant, which we write as K_m.

Figures

In the public version of the text, certain copyrighted figures are missing. The filenames for these figures are listed and many of the figures can be looked up in the following references:

- `Cou08` - *Mechanisms in Transcriptional Regulation* by A. J. Courey [17]
- `GNM93` - J. Greenblatt, J. R. Nodwell and S. W. Mason, "Transcriptional antitermination" [34]
- `Mad07` - *From a to alpha: Yeast as a Model for Cellular Differentiation* by H. Madhani [58]
- `MBoC` - *The Molecular Biology of the Cell* by Alberts et al. [2]
- `PKT08` - *Physical Biology of the Cell* by Phillips, Kondev and Theriot [72]

The remainder of the filename lists the chapter and figure number.

Review

Comments intended for reviewers are marked as in this paragraph. These comments generally explain missing material that will be included in the final text.

Chapter 1
Introductory Concepts

This chapter provides a brief introduction to concepts from systems biology, tools from differential equations and control theory, and approaches to modeling, analysis and design of biomolecular feedback systems. We begin with a discussion of the role of modeling, analysis and feedback in biological systems. This is followed by a short review of key concepts and tools from control and dynamical systems theory, intended to provide insight into the main methodology described in the text. Finally, we give a brief introduction to the field of synthetic biology, which is the primary topic of the latter portion of the text. Readers who are familiar with one or more of these areas can skip the corresponding sections without loss of continuity.

1.1 Systems Biology: Modeling, Analysis and Role of Feedback

At a variety of levels of organization—from molecular to cellular to organismal—biology is becoming more accessible to approaches that are commonly used in engineering: mathematical modeling, systems theory, computation and abstract approaches to synthesis. Conversely, the accelerating pace of discovery in biological science is suggesting new design principles that may have important practical applications in human-made systems. This synergy at the interface of biology and engineering offers many opportunities to meet challenges in both areas. The guiding principles of feedback and control are central to many of the key questions in biological science and engineering and can play an enabling role in understanding the complexity of biological systems.

In this section we summarize our view on the role that modeling and analysis should (eventually) play in the study and understanding of biological systems, and discuss some of the ways in which an understanding of feedback principles in biology can help us better understand and design complex biomolecular circuits.

There are a wide variety of biological phenomena that provide a rich source of examples for control, including gene regulation and signal transduction; hormonal, immunological, and cardiovascular feedback mechanisms; muscular control and locomotion; active sensing, vision, and proprioception; attention and consciousness; and population dynamics and epidemics. Each of these (and many more) provide opportunities to figure out what works, how it works, and what can be done to affect it. Our focus here is at the molecular scale, but the principles and approach that we describe can also be applied at larger time and length scales.

Modeling and analysis

Over the past several decades, there have been significant advances in modeling capabilities for biological systems that have provided new insights into the complex interactions of the molecular-scale processes that implement life. Reduced-order modeling has become commonplace as a mechanism for describing and documenting experimental results and high-dimensional stochastic models can now be simulated in reasonable periods of time to explore underlying stochastic effects. Coupled with our ability to collect large amounts of data from flow cytometry, micro-array analysis, single-cell microscopy, and other modern experimental techniques, our understanding of biomolecular processes is advancing at a rapid pace.

Unfortunately, although models are becoming much more common in biological studies, they are still far from playing the central role in explaining complex biological phenomena. Although there are exceptions, the predominant use of models is to "document" experimental results: a hypothesis is proposed and tested using careful experiments, and then a model is developed to match the experimental results and help demonstrate that the proposed mechanisms can lead to the observed behavior. This necessarily limits our ability to explain complex phenomena to those for which controlled experimental evidence of the desired phenomena can be obtained.

This situation is much different than standard practice in the physical sciences and engineering, as illustrated in Figure 1.1 (in the context of modeling, analysis, and control design for gas turbine aeroengines). In those disciplines, experiments are routinely used to help build models for individual components at a variety of levels of detail, and then these component-level models are interconnected to obtain a system-level model. This system-level model, carefully built to capture the appropriate level of detail for a given question or hypothesis, is used to explain, predict, and systematically analyze the behaviors of a system. Because of the ways in which models are viewed, it becomes possible to prove (or invalidate) a hypothesis through analysis of the model, and the fidelity of the models is such that decisions can be made based on them. Indeed, in many areas of modern engineering—including electronics, aeronautics, robotics, and chemical processing, to name a few—models play a primary role in the understanding of the underlying physics and/or chemistry, and these models are used in predictive ways to explore design tradeoffs and failure scenarios.

A key element in the successful application of modeling in engineering disciplines is the use of *reduced-order models* that capture the underlying dynamics of the system without necessarily modeling every detail of the underlying mechanisms. These reduced order models are often coupled with schematics diagrams, such as those shown in Figure 1.2, to provide a high level view of a complex system. The generation of these reduced-order models, either directly from data or through analytical or computational methods, is critical in the effective applica-

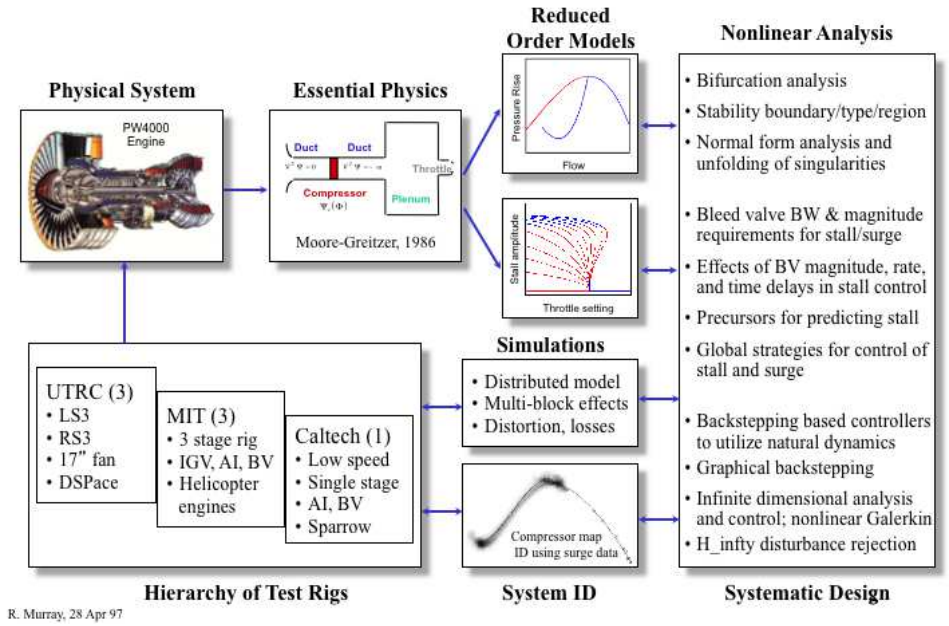


Figure 1.1: Sample modeling, analysis and design framework for an engineering system.

tion of modeling since modeling of the detailed mechanisms produces high fidelity models that are too complicated to use with existing tools for analysis and design. One area in which the development of reduced order models is fairly advanced is in control theory, where input/output models, such as block diagrams and transfer functions are used to capture structured representations of dynamics at the appropriate level of fidelity for the task at hand [1].

While developing predictive models and corresponding analysis tools for biology is much more difficult, it is perhaps even more important that biology make use of models, particularly reduced-order models, as a central element of understanding. Biological systems are by their nature extremely complex and can behave in counterintuitive ways. Only by capturing the many interacting aspects of the system in a formal model can we ensure that we are reasoning properly about its behavior, especially in the presence of uncertainty. To do this will require substantial effort in building models that capture the relevant dynamics at the proper scales (depending on the question being asked) as well as building an analytical framework for answering questions of biological relevance.

The good news is that a variety of new techniques, ranging from experiments to computation to theory, are enabling us to explore new approaches to modeling that attempt to address some of these challenges. In this text we focus on the use of relevant classes of reduced-order models that can be used to capture many phenomena of biological relevance.

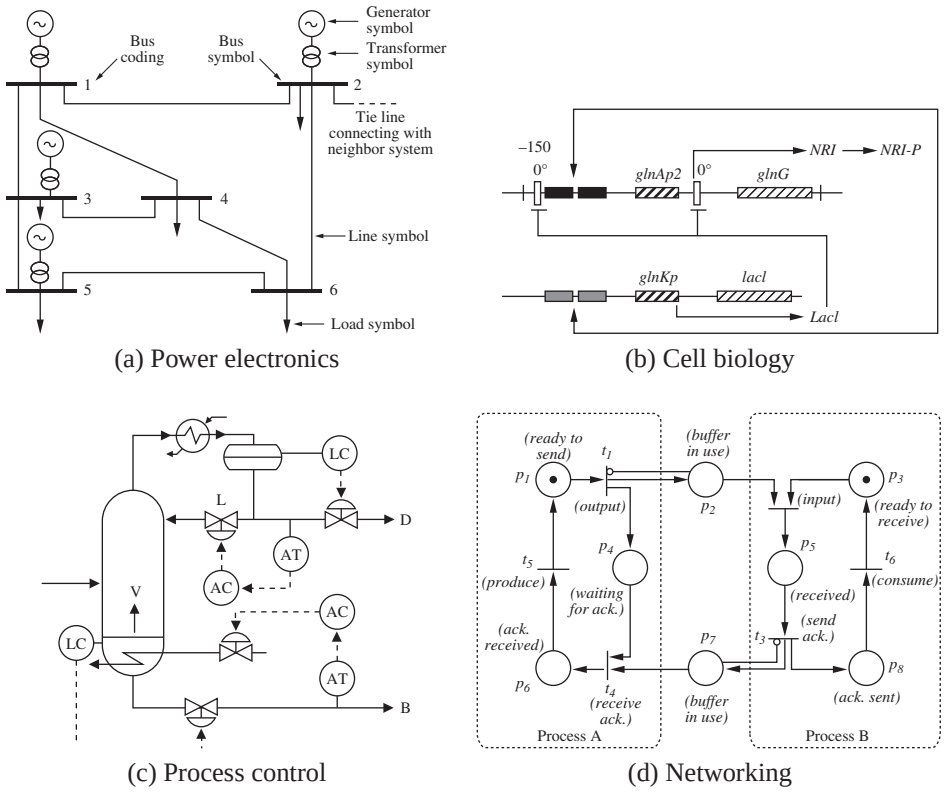


Figure 1.2: Schematic diagrams representing models in different disciplines. Each diagram is used to illustrate the dynamics of a feedback system: (a) electrical schematics for a power system [53], (b) a biological circuit diagram for a synthetic clock circuit [5], (c) a process diagram for a distillation column [80] and (d) a Petri net description of a communication protocol.

Dynamic behavior and phenotype

One of the key needs in developing a more systematic approach to the use of models in biology is to become more rigorous about the various behaviors that are important for biological systems. One of the key concepts that needs to be formalized is the notion of "phenotype". This term is often associated with the existence of an equilibrium point in a reduced-order model for a system, but clearly more complex (non-equilibrium) behaviors can occur and the "phenotypic response" of a system to an input may not be well-modeled by a steady operating condition. Even more problematic is determining which regulatory structures are "active" in a given phenotype (versus those for which there is a regulatory pathway that is saturated and hence not active).

Figure 1.3 shows a graphical representation of a class of systems that captures

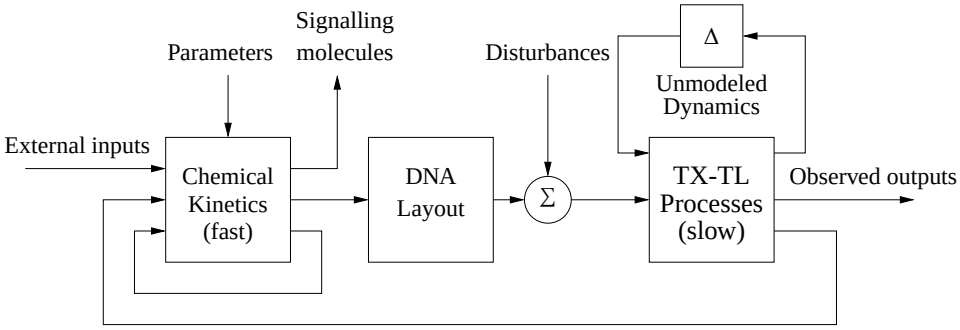


Figure 1.3: Conceptual modeling framework for biomolecular feedback systems. The chemical kinetics block represents reactions between molecular species, resulting in signaling molecules and bound promoters. The DNA layout block accounts for the organization of the DNA, which may be "rewired" to achieve a desired function. The TX-TL processes block represents the core transcription and translation processes, which are often much slower than the reactions between various species. The inputs and outputs of the various blocks represent interconnections and external interactions.

many of the features we are interested in. The chemical kinetics of the system are typically modeled using mass action kinetics (reaction rate equations) and represent the fast dynamics of chemical reactions. The reactions include the binding of activators and repressors to DNA, as well as the initiation of transcription. The DNA layout block represents the physical layout of the DNA, which determines which genes are controlled by which promoters. The core processes of transcription (TX) and translation (TL) represent the slow dynamics (relative to the chemical kinetics) of protein expression (including maturation).

Several other inputs and outputs are represented in the figure. In the chemical kinetics block, we allow external inputs, such as chemical inducers, and external parameters (rate parameters, enzyme concentrations, etc) that will effect the reactions that we are trying to capture in our model. We also include a (simplified) notion of disturbances, represented in the diagram as an external input that affects the rate of transcription. This disturbance is typically a stochastic input that represents the fact that gene expression can be noisy. In terms of outputs, we capture two possibilities in the diagram: small molecule outputs—often used for signaling to other subsystems but which could include outputs from metabolic processes—and protein outputs, such as as fluorescent reporters.

Another feature of the diagram is the block labeled "unmodeled dynamics", which represents the fact that our models of the core processes of gene expression are likely to be simplified models that ignore many details. These dynamics are modeled as a feedback interconnection with transcription and translation, which turns out to provide a rich framework for application of tools from control theory (but unfortunately one that we will not explore in great detail within this text). Tools for understanding this class of uncertainty are available for both linear and

nonlinear control systems [1] and allow stability and performance analyses in the presence of uncertainty.

The combination of partially unknown parameters, external disturbances, and unmodeled dynamics are collectively referred to as *model uncertainty* and are an important element of our analysis of biomolecular feedback systems. Often we will analyze the dynamic behavior of a system assuming that the parameters are known, disturbances are small and our models are accurate. This analysis can give valuable insights into the behavior of the system, but it is important to make sure that this behavior is robust with respect to uncertainly, a topic that we will discuss in some detail in Chapter 3.

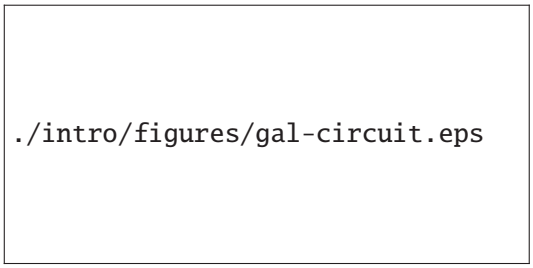
A somewhat common situation is that a system may have multiple equilibrium points and the "phenotype" of the system is represented by the particular equilibrium point that the system converges to. In the simplest case, we can have *bistability*, in which there are two equilibrium points x_{1e} and x_{2e} for a fixed set of parameters. Depending on the initial conditions and external inputs, a given system may end up near one equilibrium point or the other, providing two distinct phenotypes. A model with bistability (or multi-stability) provides one method of modeling memory in a system: the cell or organism remembers its history by virtue of the equilibrium point to which it has converted.

For more complex phenotypes, where the subsystems are not at a steady operating point, one can consider temporal patterns such as limit cycles (periodic orbits) or non-equilibrium input/output responses. Analysis of these more complicated behaviors requires more sophisticated tools, but again model-based analysis of stability and input/output responses can be used to characterize the phenotypic behavior of a biological system under different conditions or contexts.

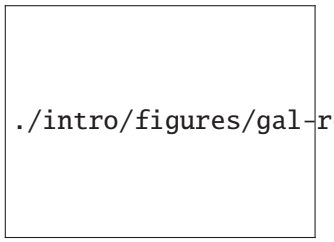
Additional types of analysis that can be applied to systems of this form include sensitivity analysis (dependence of solution properties on selected parameters), uncertainty analysis (impact of disturbances, unknown parameters and unmodeled dynamics), bifurcation analysis (changes in phenotype as a function of input levels, context or parameters) and probabilistic analysis (distributions of states as a function of distributions of parameters, initial conditions or inputs). In each of these cases, there is a need to extend existing tools to exploit the particular structure of the problems we consider, as well as modify the techniques to provide relevance to biological questions.

Stochastic behavior

Another important feature of many biological systems is stochasticity: biological responses have an element of randomness so that even under carefully control conditions, the response of a system to a given input may vary from experiment to experiment. This randomness can have many possible sources, including external perturbations that are modeled as stochastic processes and internal processes such



(a) Galactose control network



(b) Pathway response

Figure 1.4: Galactose response in yeast [93]. (a) GAL signaling circuitry showing a number of different feedback pathways that are used to detect the presence of galactose and switch on the metabolic pathway. (b) Pathway activity as a function of galactose concentration. The points at each galactose concentration represent the activity level of the galactose metabolic pathway in an individual cell. Black dots indicate the mean of a Gaussian mixture model classification [92]. Small random deviations were added to each galactose concentration (horizontal axis) to better visualize the distributions.

as molecular binding and unbinding, whose stochasticity stems from the underlying thermodynamics of molecular reactions.

While for many engineered systems it is common to try to eliminate stochastic behavior (yielding a "deterministic" response), for biological systems there appear to be many situations in which stochasticity is important for the way in which organisms survive. In biology, nothing is 100% and so there is always some chance that two identical organisms will respond differently. Thus viruses are never completely contagious and so some organisms will survive, and DNA replication is never error free, and so mutations and evolution can occur. In studying circuits where these types of effects are present, it thus becomes important to study the distribution of responses of a given biomolecular circuit, and to collect data in a manner that allows us to quantify these distributions.

One important indication of stochastic behavior is *bimodality*. We say that a circuit or system is bimodal if the response of the system to a given input or condition has two or more distinguishable classes of behaviors. An example of bimodality is shown in Figure 1.4, which shows the response of the galactose metabolic machinery in yeast. We see from the figure that even though genetically identical organisms are exposed to the same external environment (a fixed galactose concentration), the amount of activity in individual cells can have a large amount of variability. At some concentrations there are clearly two subpopulations of cells: those in which the galactose metabolic pathway is turned on (higher reporter fluorescence values on the y axis) and those for which it is off (lower reporter fluorescence).

Another characterization of stochasticity in cells is the separation of noisiness in protein expression into two categories: "intrinsic" noise and "extrinsic" noise.

Roughly speaking, extrinsic noise represents variability in gene expression that effects all proteins in the cell in a correlated way. Extrinsic noise can be due to environmental changes that affect the entire cell (temperature, pH, oxygen level) or global changes in internal factors such as energy or metabolite levels (perhaps due to metabolic loading). Intrinsic noise, on the other hand, is the variability due to the inherent randomness of molecular events inside the cell and represents a collection of independent random processes. One way to attempt to measure the amount of intrinsic and extrinsic noise is to take two identical copies of a biomolecular circuit and compare their responses [24, 88]. Correlated variations in the output of the circuits corresponds (roughly) to extrinsic noise and uncorrelated variations to intrinsic noise [40, 88].

The types of models that are used to capture stochastic behavior are very different than those used for deterministic responses. Instead of writing differential equations that track average concentration levels, we must keep track of the individual events that can occur with some probability per unit time (or "propensity"). We will explore the methods for modeling and analysis of stochastic systems in Chapter 4.

1.2 The Cell as a System

The molecular processes inside a cell determine its behavior and are responsible for metabolizing nutrients, generating motion, enabling procreation and carrying out the other functions of the organism. In multi-cellular organisms, different types of cells work together to enable more complex functions. In this section we briefly describe the role of dynamics and control within a cell and discuss the basic processes that govern its behavior and its interactions with its environment. We assume knowledge of the basics of cell biology at the level found in standard textbooks on cell biology such as Alberts *et al.* [2] or Phillips *et al.* [72].

Figure 1.5 shows a schematic of the major components in the cell: sensing, signaling, regulation, and metabolism. Sensing of environmental signals typically occurs through membrane receptors that are specific to different molecules. Cells can also respond to light or pressure, allowing the cell to sense the environment, including other cells. There are several types of receptors, some allow the signaling molecules in the environment to enter the cell wall, such as in the case of ion channels. Others activate proteins on the internal part of the cell membrane once they externally bind to the signaling molecule, such as enzyme-linked receptors or G-protein coupled receptors.

As a consequence of the sensing, a cascade of signal transduction occurs (signaling), in which proteins are sequentially activated by (usually) receiving phosphate groups from ATP molecules through the processes of phosphorylation and/or phosphotransfer. These cascades transmit information to downstream processes, such as gene expression, by amplifying the information and dynamically filtering

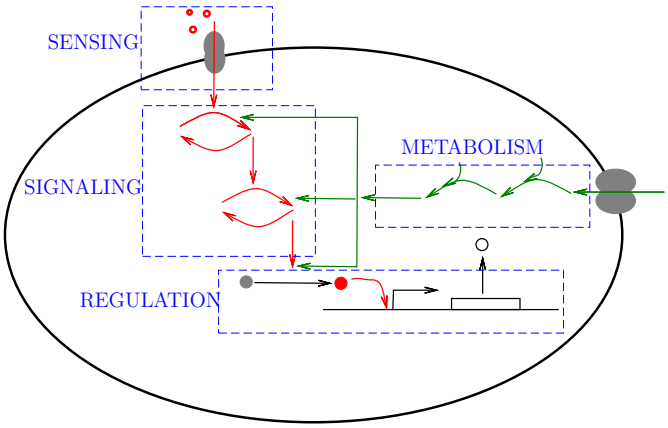


Figure 1.5: The cell as a system. The major subsystems are sensing, signaling, regulation, and metabolism.

signals to select for useful features. The temporal dynamics of environmental signals and the kinetic properties of the stages in the signaling cascades determine how a signal is transmitted/filtered. At the bottom stages of signaling cascades, proteins are activated to become transcription factors, which can activate or repress the expression of other proteins through regulation of gene expression. The temporal dynamics of this regulation, with timescales in the range of minutes to hours, are usually much slower than that of the transmission in the signaling pathway, which has timescales ranging from subseconds to seconds. "Orthogonally" to signaling cascades, metabolic pathways, such as the glycolysis pathway, are in charge of producing the necessary resources for all the other processes in the cells. Through these pathways, nutrients in the environment, such as glucose, are broken down through a series of enzymatic reactions, producing, among other products, ATP, which is the energy currency in the cell used for many of the reactions, including those involved in signaling and gene expression.

Example: Chemotaxis

As an example of a sensing-transmission-actuation process in the cell, we consider *chemotaxis*, the process by which micro-organisms move in response to chemical stimuli. Examples of chemotaxis include the ability of organisms to move in the direction of nutrients or move away from toxins in the environment. Chemotaxis is called *positive chemotaxis* if the motion is in the direction of the stimulus and *negative chemotaxis* if the motion is away from the stimulant.

The chemotaxis system in *E. coli* consists of a sensing system that detects the presence of nutrients, an actuation system that propels the organism in its environment, and control circuitry that determines how the cell should move in the

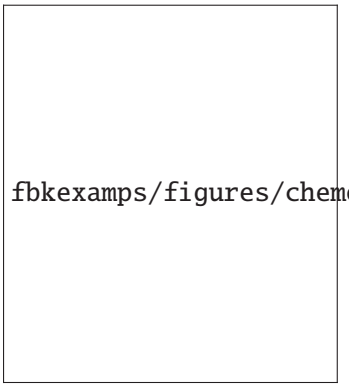


Figure 1.6: A simplified circuit diagram for chemotaxis, showing the biomolecular processes involved in regulating flagellar motion.

presence of chemicals that stimulate the sensing system. The approximate location of these elements are shown in Figure 1.6. The sensing component is responsible for detecting the presence of ligands in the environment and initiating signaling cascades. The computation component, also realized through protein phosphorylation, implements a feedback (integral) controller that allows the bacterium to adapt to changes in the environmental ligand concentration. This adaptation occurs by an actuator that allows the bacterium to ultimately move in the direction in which the ligand concentration increases.

The actuation system in the *E. coli* consists of a set of flagella that can be spun using a flagellar motor embedded in the outer membrane of the cell, as shown in Figure 1.7a. When the flagella all spin in the counter clockwise direction, the individual flagella form a bundle and cause the organism to move roughly in a straight line. This behavior is called a "run" motion. Alternatively, if the flagella spin in the clockwise direction, the individual flagella do not form a bundle and the organism "tumbles", causing it to rotate (Figure 1.7b). The selection of the motor direction is controlled by the protein CheY: if phosphorylated CheY binds to the motor complex, the motor spins clockwise (tumble), otherwise it spins counter-clockwise (run). As a consequence, the chemotaxis mechanisms is stochastic in nature, with biased random motions causing the average behavior to be either positive, negative or neutral (in the absence of stimuli).

1.3 Control and Dynamical Systems Tools[1]

To study the complex dynamics and feedback present in biological systems, we will make use of mathematical models combined with analytical and computational tools. In this section we present a brief introduction to some of the key concepts

[1]The material in this section is adapted from *Feedback Systems*, Chapter 1 [1].

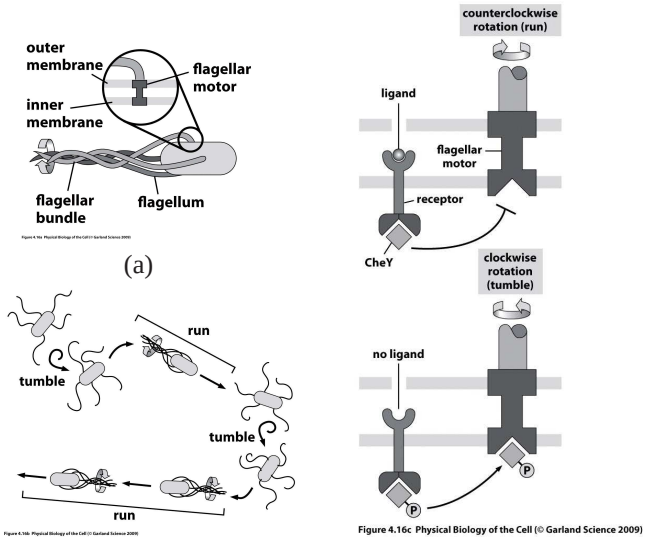


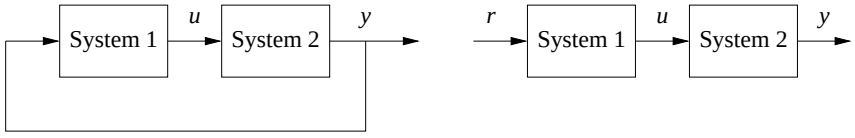
Figure 1.7: Bacterial chemotaxis. Figures from Phillips, Kondev and Theriot [72]; used with permission of Garland Science.

from control and dynamical systems that are relevant for the study of biomolecular systems considered in later chapters. More details on the application of specific concepts listed here to biomolecular systems is provided in the main body of the text. Readers who are familiar with introductory concepts in dynamical systems and control, at the level described in Åström and Murray [1] for example, can skip this section.

Dynamics, feedback and control

A *dynamical system* is a system whose behavior changes over time, often in response to external stimulation or forcing. The term *feedback* refers to a situation in which two (or more) dynamical systems are connected together such that each system influences the other and their dynamics are thus strongly coupled. Simple causal reasoning about a feedback system is difficult because the first system influences the second and the second system influences the first, leading to a circular argument. This makes reasoning based on cause and effect tricky, and it is necessary to analyze the system as a whole. A consequence of this is that the behavior of feedback systems is often counterintuitive, and it is therefore often necessary to resort to formal methods to understand them.

Figure 1.8 illustrates in block diagram form the idea of feedback. We often use the terms *open loop* and *closed loop* when referring to such systems. A system is said to be a closed loop system if the systems are interconnected in a cycle, as



(a) Closed loop (b) Open loop

Figure 1.8: Open and closed loop systems. (a) The output of system 1 is used as the input of system 2, and the output of system 2 becomes the input of system 1, creating a closed loop system. (b) The interconnection between system 2 and system 1 is removed, and the system is said to be open loop.

shown in Figure 1.8a. If we break the interconnection, we refer to the configuration as an open loop system, as shown in Figure 1.8b.

Biological systems make use of feedback in an extraordinary number of ways, on scales ranging from molecules to cells to organisms to ecosystems. One example is the regulation of glucose in the bloodstream through the production of insulin and glucagon by the pancreas. The body attempts to maintain a constant concentration of glucose, which is used by the body's cells to produce energy. When glucose levels rise (after eating a meal, for example), the hormone insulin is released and causes the body to store excess glucose in the liver. When glucose levels are low, the pancreas secretes the hormone glucagon, which has the opposite effect. Referring to Figure 1.8, we can view the liver as system 1 and the pancreas as system 2. The output from the liver is the glucose concentration in the blood, and the output from the pancreas is the amount of insulin or glucagon produced. The interplay between insulin and glucagon secretions throughout the day helps to keep the blood-glucose concentration constant, at about 90 mg per 100 mL of blood.

Feedback has many interesting properties that can be exploited in designing systems. As in the case of glucose regulation, feedback can make a system resilient toward external influences. It can also be used to create linear behavior out of nonlinear components, a common approach in electronics. More generally, feedback allows a system to be insensitive both to external disturbances and to variations in its individual elements.

Feedback has potential disadvantages as well. It can create dynamic instabilities in a system, causing oscillations or even runaway behavior. Another drawback, especially in engineering systems, is that feedback can introduce unwanted sensor noise into the system, requiring careful filtering of signals. It is for these reasons that a substantial portion of the study of feedback systems is devoted to developing an understanding of dynamics and a mastery of techniques in dynamical systems.

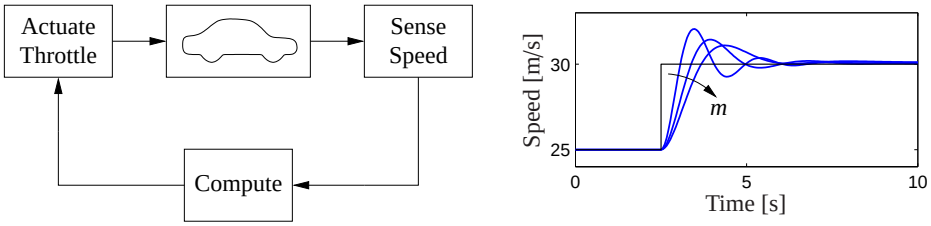


Figure 1.9: A feedback system for controlling the speed of a vehicle. In the block diagram on the left, the speed of the vehicle is measured and compared to the desired speed within the "Compute" block. Based on the difference in the actual and desired speeds, the throttle (or brake) is used to modify the force applied to the vehicle by the engine, drivetrain and wheels. The figure on the right shows the response of the control system to a commanded change in speed from 25 m/s to 30 m/s. The three different curves correspond to differing masses of the vehicle, between 1000 and 3000 kg, demonstrating the robustness of the closed loop system to a very large change in the vehicle characteristics.

Feedback properties

Feedback is a powerful idea that is used extensively in natural and technological systems. The principle of feedback is simple: implement correcting actions based on the difference between desired and actual performance. In engineering, feedback has been rediscovered and patented many times in many different contexts. The use of feedback has often resulted in vast improvements in system capability, and these improvements have sometimes been revolutionary, as discussed above. The reason for this is that feedback has some truly remarkable properties, which we discuss briefly here.

Robustness to Uncertainty. One of the key uses of feedback is to provide robustness to uncertainty. By measuring the difference between the sensed value of a regulated signal and its desired value, we can supply a corrective action. If the system undergoes some change that affects the regulated signal, then we sense this change and try to force the system back to the desired operating point.

As an example of this principle, consider the simple feedback system shown in Figure 1.9. In this system, the speed of a vehicle is controlled by adjusting the amount of gas flowing to the engine. Simple *proportional-integral* (PI) feedback is used to make the amount of gas depend on both the error between the current and the desired speed and the integral of that error. The plot on the right shows the results of this feedback for a step change in the desired speed and a variety of different masses for the car, which might result from having a different number of passengers or towing a trailer. Notice that independent of the mass (which varies by a factor of 3!), the steady-state speed of the vehicle always approaches the desired speed and achieves that speed within approximately 5 s. Thus the performance of the system is robust with respect to this uncertainty.

Another early example of the use of feedback to provide robustness is the neg-

ative feedback amplifier. When telephone communications were developed, amplifiers were used to compensate for signal attenuation in long lines. A vacuum tube was a component that could be used to build amplifiers. Distortion caused by the nonlinear characteristics of the tube amplifier together with amplifier drift were obstacles that prevented the development of line amplifiers for a long time. A major breakthrough was the invention of the feedback amplifier in 1927 by Harold S. Black, an electrical engineer at Bell Telephone Laboratories. Black used *negative feedback*, which reduces the gain but makes the amplifier insensitive to variations in tube characteristics. This invention made it possible to build stable amplifiers with linear characteristics despite the nonlinearities of the vacuum tube amplifier.

Feedback is also pervasive in biological systems, where transcriptional, translational and allosteric mechanisms are used to regulate internal concentrations of various species, and much more complex feedbacks are used to regulate properties at the organism level (such as body temperature, blood pressure and circadian rhythm). One difference in biological systems is that the separation of sensing, actuation and computation, a common approach in most engineering control systems, is less evident. Instead, the dynamics of the molecules that sense the environmental condition and make changes to the operation of internal components may be integrated together in ways that make it difficult to untangle the operation of the system. Similarly, the "reference value" to which we wish to regulate a system may not be an explicit signal, but rather a consequence of many different changes in the dynamics that are coupled back to the regulatory elements. Hence we do not see a clear "setpoint" for the desired ATP concentration, blood oxygen level or body temperature, for example. These difficulties complicate our analysis of biological systems, though many important insights can still be obtained.

Design of Dynamics. Another use of feedback is to change the dynamics of a system. Through feedback, we can alter the behavior of a system to meet the needs of an application: systems that are unstable can be stabilized, systems that are sluggish can be made responsive and systems that have drifting operating points can be held constant. Control theory provides a rich collection of techniques to analyze the stability and dynamic response of complex systems and to place bounds on the behavior of such systems by analyzing the gains of linear and nonlinear operators that describe their components.

An example of the use of control in the design of dynamics comes from the area of flight control. The following quote, from a lecture presented by Wilbur Wright to the Western Society of Engineers in 1901 [64], illustrates the role of control in the development of the airplane:

> Men already know how to construct wings or airplanes, which when driven through the air at sufficient speed, will not only sustain the weight of the wings themselves, but also that of the engine, and of the engineer as well. Men also know how to build engines and screws

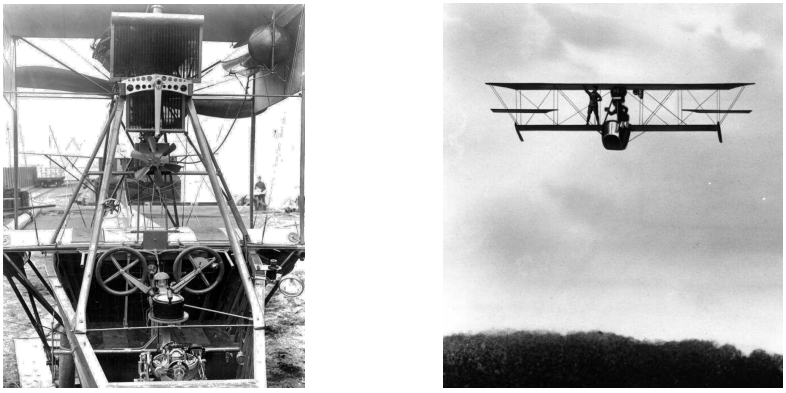


Figure 1.10: Aircraft autopilot system. The Sperry autopilot (left) contained a set of four gyros coupled to a set of air valves that controlled the wing surfaces. The 1912 Curtiss used an autopilot to stabilize the roll, pitch and yaw of the aircraft and was able to maintain level flight as a mechanic walked on the wing (right) [42].

> of sufficient lightness and power to drive these planes at sustaining speed ... Inability to balance and steer still confronts students of the flying problem ... When this one feature has been worked out, the age of flying will have arrived, for all other difficulties are of minor importance.

The Wright brothers thus realized that control was a key issue to enable flight. They resolved the compromise between stability and maneuverability by building an airplane, the Wright Flyer, that was unstable but maneuverable. The Flyer had a rudder in the front of the airplane, which made the plane very maneuverable. A disadvantage was the necessity for the pilot to keep adjusting the rudder to fly the plane: if the pilot let go of the stick, the plane would crash. Other early aviators tried to build stable airplanes. These would have been easier to fly, but because of their poor maneuverability they could not be brought up into the air. By using their insight and skillful experiments the Wright brothers made the first successful flight at Kitty Hawk in 1903.

Since it was quite tiresome to fly an unstable aircraft, there was strong motivation to find a mechanism that would stabilize an aircraft. Such a device, invented by Sperry, was based on the concept of feedback. Sperry used a gyro-stabilized pendulum to provide an indication of the vertical. He then arranged a feedback mechanism that would pull the stick to make the plane go up if it was pointing down, and vice versa. The Sperry autopilot was the first use of feedback in aeronautical engineering, and Sperry won a prize in a competition for the safest airplane in Paris in 1914. Figure 1.10 shows the Curtiss seaplane and the Sperry autopilot. The autopilot is a good example of how feedback can be used to stabilize an unstable system and hence “design the dynamics” of the aircraft.

One of the other advantages of designing the dynamics of a device is that it allows for increased modularity in the overall system design. By using feedback to create a system whose response matches a desired profile, we can hide the complexity and variability that may be present inside a subsystem. This allows us to create more complex systems by not having to simultaneously tune the responses of a large number of interacting components. This was one of the advantages of Black's use of negative feedback in vacuum tube amplifiers: the resulting device had a well-defined linear input/output response that did not depend on the individual characteristics of the vacuum tubes being used.

Drawbacks of Feedback. While feedback has many advantages, it also has some drawbacks. Chief among these is the possibility of instability if the system is not designed properly. We are all familiar with the undesirable effects of feedback when the amplification on a microphone is turned up too high in a room. This is an example of feedback instability, something that we obviously want to avoid. This is tricky because we must design the system not only to be stable under nominal conditions but also to remain stable under all possible perturbations of the dynamics. In biomolecular systems, these types of instabilities may exhibit themselves as situations in which cells no longer function properly due to over expression of engineered genetic components, or small fluctuations in parameters cause the system to suddenly cease to function properly.

In addition to the potential for instability, feedback inherently couples different parts of a system. One common problem is that feedback often injects "crosstalk" into the system. By coupling different parts of a biomolecular circuit, the fluctuations in one part of the circuit affect other parts, which themselves may couple to the initial source of the fluctuations. If we are designing a biomolecular system, this crosstalk may make affect our ability to design independent "modules" whose behavior can described in isolation.

Coupled to the problem of crosstalk is the substantial increase in complexity that results when embedding multiple feedback loops in a system. An early engineering example of this was the use of microprocessor-based feedback systems in automobiles. The use of microprocessors in automotive applications began in the early 1970s and was driven by increasingly strict emissions standards, which could be met only through electronic controls. Early systems were expensive and failed more often than desired, leading to frequent customer dissatisfaction. It was only through aggressive improvements in technology that the performance, reliability and cost of these systems allowed them to be used in a transparent fashion. Even today, the complexity of these systems is such that it is difficult for an individual car owner to fix problems. While nature has evolved many feedback structures that are robust and reliable, engineered biomolecular systems are still quite rudimentary and we can anticipate that as we increase the use of feedback to compensate for uncertainty, we will see a similar period in which engineers must overcome a steep learning curve before we can get robust and reliable behavior as a matter of

course.

Feedforward. Feedback is reactive: there must be an error before corrective actions are taken. However, in some circumstances it is possible to measure a disturbance before it enters the system, and this information can then be used to take corrective action before the disturbance has influenced the system. The effect of the disturbance is thus reduced by measuring it and generating a control signal that counteracts it. This way of controlling a system is called *feedforward*. Feedforward is particularly useful in shaping the response to command signals because command signals are always available. Since feedforward attempts to match two signals, it requires good process models; otherwise the corrections may have the wrong size or may be badly timed.

The ideas of feedback and feedforward are very general and appear in many different fields. In economics, feedback and feedforward are analogous to a market-based economy versus a planned economy. In business, a feedforward strategy corresponds to running a company based on extensive strategic planning, while a feedback strategy corresponds to a reactive approach. In biology, feedforward has been suggested as an essential element for motion control in humans that is tuned during training. Experience indicates that it is often advantageous to combine feedback and feedforward, and the correct balance requires insight and understanding of their respective properties.

Positive Feedback. In most of control theory, the emphasis is on the role of *negative feedback*, in which we attempt to regulate the system by reacting to disturbances in a way that decreases the effect of those disturbances. In some systems, particularly biological systems, *positive feedback* can play an important role. In a system with positive feedback, the increase in some variable or signal leads to a situation in which that quantity is further increased through its dynamics. This has a destabilizing effect and is usually accompanied by a saturation that limits the growth of the quantity. Although often considered undesirable, this behavior is used in biological (and engineering) systems to obtain a very fast response to a condition or signal.

One example of the use of positive feedback is to create switching behavior, in which a system maintains a given state until some input crosses a threshold. Hysteresis is often present so that noisy inputs near the threshold do not cause the system to jitter. This type of behavior is called *bistability* and is often associated with memory devices.

1.4 Input/Output Modeling[2]

A model is a mathematical representation of a physical, biological or information system. Models allow us to reason about a system and make predictions about

[2]The material in this section is adapted from *Feedback Systems*, Sections 2.1–2.2 [1].

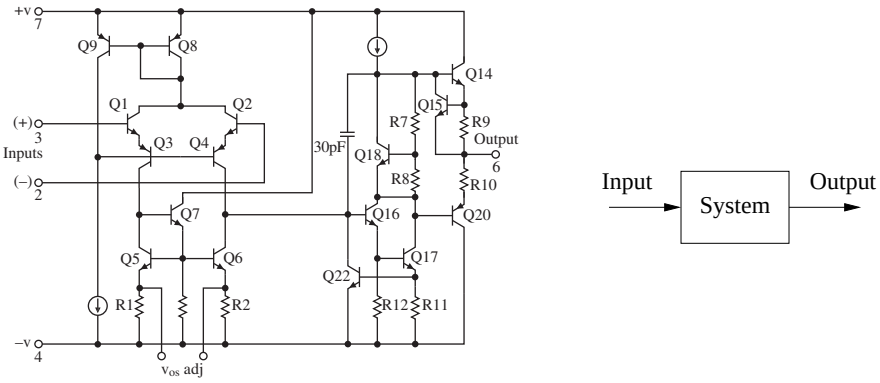


Figure 1.11: Illustration of the input/output view of a dynamical system. The figure on the left shows a detailed circuit diagram for an electronic amplifier; the one on the right is its representation as a block diagram.

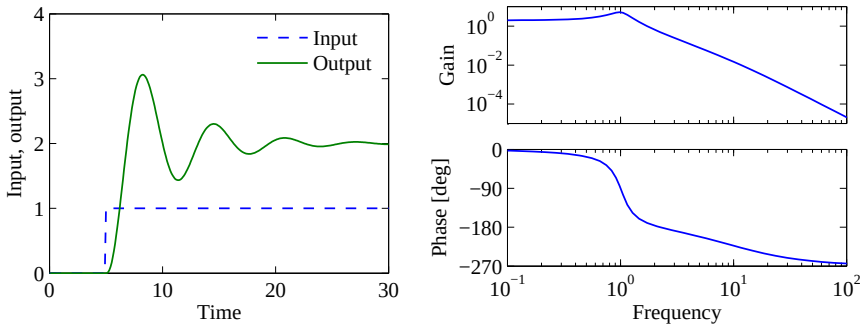
how a system will behave. In this text, we will mainly be interested in models of dynamical systems describing the input/output behavior of systems, and we will often work in "state space" form. In the remainder of this section we provide an overview of some of the key concepts in input/output modeling. The mathematical details introduced here are explored more fully in Chapter 3.

The heritage of electrical engineering

The approach to modeling that we take builds on the view of models that emerged from electrical engineering, where the design of electronic amplifiers led to a focus on input/output behavior. A system was considered a device that transforms inputs to outputs, as illustrated in Figure 1.11. Conceptually an input/output model can be viewed as a giant table of inputs and outputs. Given an input signal $u(t)$ over some interval of time, the model should produce the resulting output $y(t)$.

The input/output framework is used in many engineering disciplines since it allows us to decompose a system into individual components connected through their inputs and outputs. Thus, we can take a complicated system such as a radio or a television and break it down into manageable pieces such as the receiver, demodulator, amplifier and speakers. Each of these pieces has a set of inputs and outputs and, through proper design, these components can be interconnected to form the entire system.

The input/output view is particularly useful for the special class of *linear time-invariant systems*. This term will be defined more carefully below, but roughly speaking a system is linear if the superposition (addition) of two inputs yields an output that is the sum of the outputs that would correspond to individual inputs being applied separately. A system is time-invariant if the output response for a given input does not depend on when that input is applied. While most biomolecular sys-



(a) Step response (b) Frequency response

Figure 1.12: Input/output response of a linear system. The step response (a) shows the output of the system due to an input that changes from 0 to 1 at time $t = 5$ s. The frequency response (b) shows the amplitude gain and phase change due to a sinusoidal input at different frequencies.

tems are neither linear nor time-invariant, they can often be approximated by such models, often by looking at perturbations of the system from its nominal behavior, in a fixed context.

One of the reasons that linear time-invariant systems are so prevalent in modeling of input/output systems is that a large number of tools have been developed to analyze them. One such tool is the *step response*, which describes the relationship between an input that changes from zero to a constant value abruptly (a step input) and the corresponding output. The step response is very useful in characterizing the performance of a dynamical system, and it is often used to specify the desired dynamics. A sample step response is shown in Figure 1.12a.

Another way to describe a linear time-invariant system is to represent it by its response to sinusoidal input signals. This is called the *frequency response*, and a rich, powerful theory with many concepts and strong, useful results has emerged for systems that can be described by their frequency response. The results are based on the theory of complex variables and Laplace transforms. The basic idea behind frequency response is that we can completely characterize the behavior of a system by its steady-state response to sinusoidal inputs. Roughly speaking, this is done by decomposing any arbitrary signal into a linear combination of sinusoids (e.g., by using the Fourier transform) and then using linearity to compute the output by combining the response to the individual frequencies. A sample frequency response is shown in Figure 1.12b.

The input/output view lends itself naturally to experimental determination of system dynamics, where a system is characterized by recording its response to particular inputs, e.g., a step or a set of sinusoids over a range of frequencies.

The control view

When control theory emerged as a discipline in the 1940s, the approach to dynamics was strongly influenced by the electrical engineering (input/output) view. A second wave of developments in control, starting in the late 1950s, was inspired by mechanics, where the state space perspective was used. The emergence of space flight is a typical example, where precise control of the orbit of a spacecraft is essential. These two points of view gradually merged into what is today the state space representation of input/output systems.

The development of state space models involved modifying the models from mechanics to include external actuators and sensors and utilizing more general forms of equations. In control, models often take the form

$$\frac{dx}{dt} = f(x,u), \qquad y = h(x,u), \tag{1.1}$$

where x is a vector of state variables, u is a vector of control signals and y is a vector of measurements. The term dx/dt (sometimes also written as $\dot{x}$) represents the derivative of x with respect to time, now considered a vector, and f and h are (possibly nonlinear) mappings of their arguments to vectors of the appropriate dimension.

Adding inputs and outputs has increased the richness of the classical problems and led to many new concepts. For example, it is natural to ask if possible states x can be reached with the proper choice of u (reachability) and if the measurement y contains enough information to reconstruct the state (observability). These topics are addressed in greater detail in AM08.

A final development in building the control point of view was the emergence of disturbances and model uncertainty as critical elements in the theory. The simple way of modeling disturbances as deterministic signals like steps and sinusoids has the drawback that such signals cannot be predicted precisely. A more realistic approach is to model disturbances as random signals. This viewpoint gives a natural connection between prediction and control. The dual views of input/output representations and state space representations are particularly useful when modeling uncertainty since state models are convenient to describe a nominal model but uncertainties are easier to describe using input/output models (often via a frequency response description).

An interesting observation in the design of control systems is that feedback systems can often be analyzed and designed based on comparatively simple models. The reason for this is the inherent robustness of feedback systems. However, other uses of models may require more complexity and more accuracy. One example is feedforward control strategies, where one uses a model to precompute the inputs that cause the system to respond in a certain way. Another area is system validation, where one wishes to verify that the detailed response of the system performs

as it was designed. Because of these different uses of models, it is common to use a hierarchy of models having different complexity and fidelity.

State space systems

The state of a system is a collection of variables that summarize the past of a system for the purpose of predicting the future. For a biochemical system the state is composed of the variables required to account for the current context of the cell, including the concentrations of the various species and complexes that are present. It may also include the spatial locations of the various molecules. A key issue in modeling is to decide how accurately this information has to be represented. The state variables are gathered in a vector $x \in \mathbb{R}^n$ called the *state vector*. The control variables are represented by another vector $u \in \mathbb{R}^p$, and the measured signal by the vector $y \in \mathbb{R}^q$. A system can then be represented by the differential equation

$$\frac{dx}{dt} = f(x,u), \qquad y = h(x,u), \tag{1.2}$$

where $f : \mathbb{R}^n \times \mathbb{R}^p \to \mathbb{R}^n$ and $h : \mathbb{R}^n \times \mathbb{R}^p \to \mathbb{R}^q$ are smooth mappings. We call a model of this form a *state space model*.

The dimension of the state vector is called the *order* of the system. The system (1.2) is called *time-invariant* because the functions f and h do not depend explicitly on time t; there are more general time-varying systems where the functions do depend on time. The model consists of two functions: the function f gives the rate of change of the state vector as a function of state x and control u, and the function h gives the measured values as functions of state x and control u.

A system is called a *linear* state space system if the functions f and h are linear in x and u. A linear state space system can thus be represented by

$$\frac{dx}{dt} = Ax + Bu, \qquad y = Cx + Du, \tag{1.3}$$

where A, B, C and D are constant matrices. Such a system is said to be *linear and time-invariant*, or LTI for short. The matrix A is called the *dynamics matrix*, the matrix B is called the *control matrix*, the matrix C is called the *sensor matrix* and the matrix D is called the *direct term*. Frequently systems will not have a direct term, indicating that the control signal does not influence the output directly.

1.5 From Systems to Synthetic Biology

The rapidly growing field of synthetic biology seeks to use biological principles and processes to build useful engineering devices and systems. Applications of synthetic biology range from materials production (drugs, biofuels) to biological

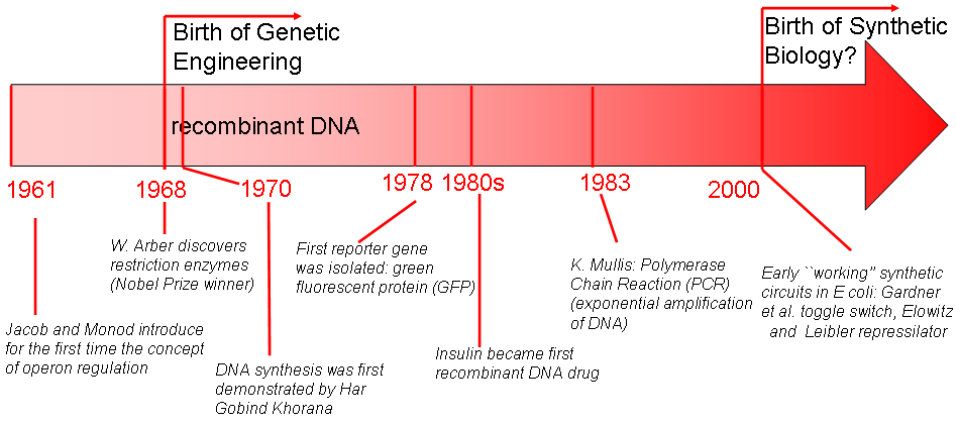


Figure 1.13: Milestones in the history of synthetic biology.

sensing and diagnostics (chemical detection, medical diagnostics) to biological machines (bioremediation, nanoscale robotics). Like many other fields at the time of their infancy (electronics, software, networks), it is not yet clear where synthetic biology will have its greatest impact. However, recent advances such as the ability to "boot up" a chemically synthesized genome [28] demonstrate the ability to synthesize systems that offer the possibility of creating devices with substantial functionality. At the same time, the tools and processes available to design systems of this complexity are much more primitive, and *de novo* synthetic circuits typically use a tiny fraction of the number of genetic elements of even the smallest microorganisms [74].

Several scientific and technological developments over the past four decades have set the stage for the design and fabrication of early synthetic biomolecular circuits (see Figure 1.13). An early milestone in the history of synthetic biology can be traced back to the discovery of mathematical logic in gene regulation. In their 1961 paper, Jacob and Monod introduced for the first time the idea of gene expression regulation through transcriptional feedback [45]. Only a few years later (1969), *restriction enzymes* that cut double-stranded DNA at specific recognition sites were discovered by Arber and co-workers [4]. These enzymes were a major enabler of recombinant DNA technology, in which genes from one organism are extracted and spliced into the chromosome of another. One of the most celebrated products of this technology was the large scale production of insulin by employing *E. coli* bacteria as a cell factory [94].

Another key innovation was the development of the polymerase chain reaction (PCR), devised in the 1980s, which allows exponential amplification of small amounts of DNA and can be used to obtain sufficient quantities for use in a variety of molecular biology laboratory protocols where higher concentrations of DNA are required. Using PCR, it is possible to "copy" genes and other DNA sequences out

of their host organisms.

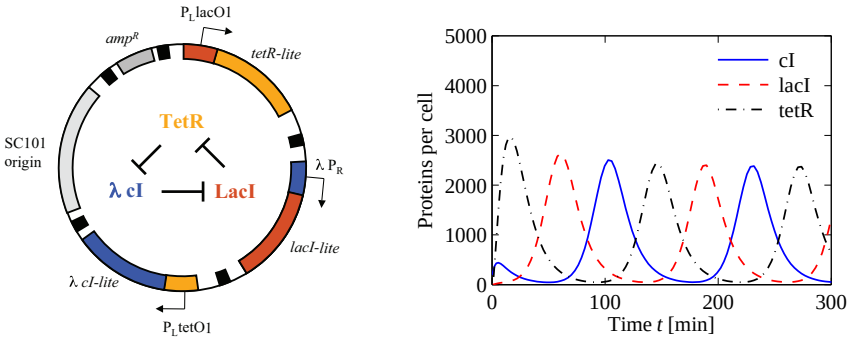
The developments of recombinant DNA technology, PCR and artificial synthesis of DNA provided the ability to "cut and paste" natural or synthetic promoters and genes in almost any fashion. This cut and paste procedure is called *cloning* and traditionally consists of four primary steps: *fragmentation*, *ligation*, *transfection* and *screening*. The DNA of interest is first isolated using restriction enzymes and/or PCR amplification. Then, a ligation procedure is employed in which the amplified fragment is inserted into a vector. The vector is often a piece of circular DNA, called a plasmid, that has been linearized by means of restriction enzymes that cleave it at appropriate restriction sites. The vector is then incubated with the fragment of interest with an enzyme called *DNA ligase*, producing a single piece of DNA with the target DNA inserted. The next step is to transfect (or transform) the DNA into living cells, where the natural replication mechanisms of the cell will duplicate the DNA when the cell divides. This process does not transfect all cells, and so a selection procedure if required to isolate those cells that have the desired DNA inserted in them. This is typically done by using a plasmid that gives the cell resistance to a specific antibiotic; cells grown in the presence of that antibiotic will only live if they contain the plasmid. Further selection can be done to insure that the inserted DNA is also present.

Once a circuit has been constructed, its performance must be verified and, if necessary, debugged. This is often done with the help of *fluorescent reporters*. The most famous of these is GFP, which was isolated from the jellyfish *Aequorea victoria* in 1978 by Shimomura [82]. Further work by Chalfie and others in the 1990s enabled the use of GFP in *E. coli* as a fluorescent reporter by inserting it into an appropriate point in an artificial circuit [16]. By using spectrofluorometry, fluorescent microscopy or flow cytometry, it is possible to measure the amount of fluorescence in individual cells or collections of cells and characterize the performance of a circuit in the presence of inducers or other factors.

Two early examples of the application of these technologies were the *repressilator* [23] and a synthetic genetic switch [27].

The repressilator is a synthetic circuit in which three proteins each repress another in a cycle. This is shown schematically in Figure 1.14a, where the three proteins are TetR, λ cI and LacI. The basic idea of the repressilator is that if TetR is present, then it represses the production of λ cI. If λ cI is absent, then LacI is produced (at the unregulated transcription rate), which in turn represses TetR. Once TetR is repressed, then λ cI is no longer repressed, and so on. If the dynamics of the circuit are designed properly, the resulting protein concentrations will oscillate, as shown in Figure 1.14b.

The repressilator can be constructed using the techniques described above. First, we can make copies of the individual promoters and genes that form our circuit by using PCR to amplify the selected sequences out of the original organisms in which they were found. TetR is the tetracycline resistance repressor protein that is found



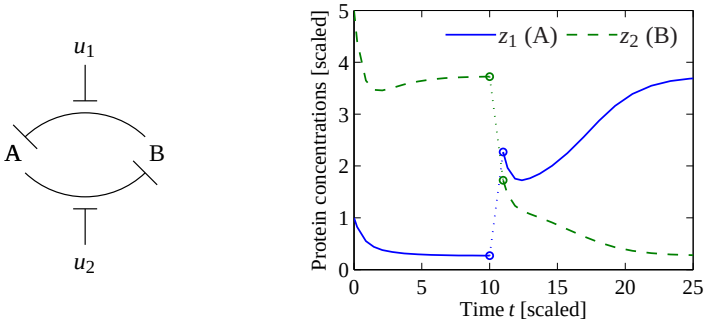
(a) Repressilator plasmid (b) Repressilator simulation

Figure 1.14: The repressilator genetic regulatory network. (a) A schematic diagram of the repressilator, showing the layout of the genes in the plasmid that holds the circuit as well as the circuit diagram (center). The flat headed arrow between the protein names represents repression. (b) A simulation of a simple model for the repressilator, showing the oscillation of the individual protein concentrations. (Figure courtesy M. Elowitz.)

in gram-negative bacteria (such as *E. coli*) and is part of the circuitry that provides resistance to tetracycline. LacI is the gene that produces *lac* repressor, responsible for turning off the *lac* operon in the lactose metabolic pathway in *E. coli* (see Section 5.1). And λ cI comes from λ phage, where it is part of the regulatory circuitry that regulates lysis and lysogeny.

By using restriction enzymes and related techniques, we can separate the natural promoters from their associated genes, and then ligate (reassemble) them in a new order and insert them into a "backbone" vector (the rest of the plasmid, including the origin of replication and appropriate antibiotic resistance). This DNA is then transformed into cells that are grown in the presence of an antibiotic, so that only those cells that contain the represilator can replicate. Finally, we can take individual cells containing our circuit and let them grow under a microscope to image fluorescent reporters coupled to the oscillator.

Another early circuit in the synthetic biology toolkit is a genetic switch built by Gardner *et al.* [27]. The genetic switch consists of two repressors connected together in a cycle, as shown in Figure 1.15a. The intuition behind this circuit is that if the gene A is being expressed, it will repress production of B and maintain its expression level (since the protein corresponding to B will not be present to repress A). Similarly, if B is being expressed, it will repress the production of A and maintain its expression level. This circuit thus implements a type of *bistability* that can be used as a simple form of memory. Figure 1.15b shows the time traces for a system, illustrating the bistable nature of the circuit. When the initial condition starts with a concentration of protein B greater than that of A, the solution converges to the equilibrium point where B is on and A is off. If A is greater than B, then the opposite situation results.



(a) Switch circuit diagram (b) Switch simulation

Figure 1.15: Stability of a genetic switch. The circuit diagram in (a) represents two proteins that are each repressing the production of the other. The inputs u_1 and u_2 interfere with this repression, allowing the circuit dynamics to be modified. The simulation in (b) shows the time response of the system starting from two different initial conditions. The initial portion of the curve corresponds to protein B having higher concentration than A, and converges to an equilibrium where A is off and B is on. At time $t = 10$, the concentrations are perturbed, moving the concentrations into a region of the state space where solutions converge to the equilibrium point with the A on and B off.

These seemingly simple circuits took years of effort to get to work, but showed that it was possible to synthesize a biological circuit that performed a desired function that was not originally present in a natural system. Today, commercial synthesis of DNA sequences and genes has become cheaper and faster, with a price often below $0.20 per base pair.[3] The combination of inexpensive synthesis technologies, new advances in cloning techniques, and improved devices for imaging and measurement has vastly simplified the process of producing a sequence of DNA that encodes a given set of genes, operator sites, promoters and other functions. These techniques are a routine part of undergraduate courses in molecular and synthetic biology.

As illustrated by the examples above, current techniques in synthetic biology have demonstrated the ability to program biological function by designing DNA sequences that implement simple circuits. Most current devices make use of transcriptional or post-transcriptional processing, resulting in very slow timescales (response times typically measured in tens of minutes to hours). This restricts their use in systems where faster response to environmental signals is needed, such as rapid detection of a chemical signal or fast response to changes in the internal environment of the cell. In addition, existing methods for biological circuit design have limited modularity (reuse of circuit elements requires substantial redesign or tuning) and typically operate in very narrow operating regimes (e.g., a single species grown in a single type of media under carefully controlled conditions). Further-

[3] As of this writing; divide by a factor of two for every two years after the publication date.

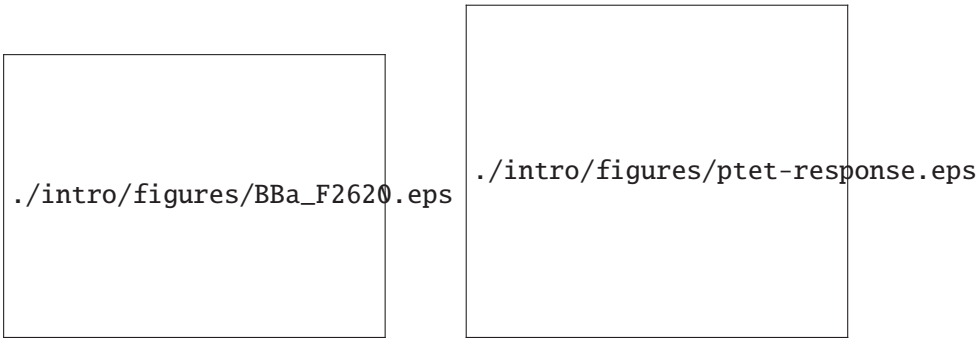


Figure 1.16: Expression of a protein using an inducible promoter [15]. (a) The circuit diagram indicates the DNA sequences that are used to construct the part (chosen from the BioBrick library). (b) The measured response of the system to a step change in the inducer level (HSL).

more, engineered circuits inserted into cells can interact with the host organism and have other unintended interactions.

As an illustration of the dynamics of synthetic devices, Figure 1.16 shows a typical response of a genetic element to an inducer molecule [15]. In this circuit, an external signal of homoserine lactone (HSL) is applied at time zero and the system reaches 10% of the steady state value in approximately 15 minutes. This response is limited in part by the time required to synthesize the output protein (GFP), including delays due to transcription, translation and folding. Since this is the response time for the underlying "actuator", circuits that are composed of feedback interconnections of such genetic elements will typically operate at 5–10 times slower speeds. While these speeds are appropriate in many applications (e.g., regulation of steady state enzyme levels for materials production), in the context of biochemical sensors or systems that must maintain a steady operating point in more rapidly changing thermal or chemical environments, this response time is too slow to be used as an effective engineering approach.

By comparison, the input/output response for the signaling component in *E. coli* chemotaxis is shown in Figure 1.17 [81]. Here the response of the kinase CheA is plotted in response to an exponential ramp in the ligand concentration. The response is extremely rapid, with the timescale measured in seconds. This rapid response is implemented by conformational changes in the proteins involved in the circuit, rather than regulation of transcription or other slower processes.

The field of synthetic biology has the opportunity to provide new approaches to solving engineering and scientific problems. Sample engineering applications include the development of synthetic circuits for producing biofuels, ultrasensitive chemical sensors, or production of materials with specific properties that are tuned to commercial needs. In addition to the potential impact on new biologically engi-

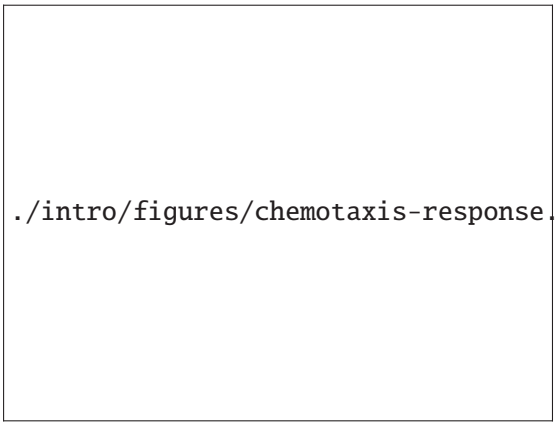


Figure 1.17: Responses of *E. coli* chemotaxis signaling network to exponential ramps in ligand concentration. Time responses of the "sensing" subsystem (from Shimizu, Tu and Berg; Molecular Systems Biology, 2010), showing the response to exponential inputs.

neered devices, there is also the potential for impact in improved understanding of biological processes. For example, many diseases such as cancer and Parkinson's disease are closely tied to kinase dysfunction. Our analysis of robust systems of kinases and the ability to synthesize systems that support or invalidate biological hypotheses may lead to a better systems understanding of failure modes that lead to such diseases.

1.6 Further Reading

There are numerous survey articles and textbooks that provide more detailed introductions to the topics introduced in this chapter. In the field of systems biology, the textbook by Alon [3] provides a broad view of some of the key elements of modern systems biology. A more comprehensive set of topics is covered in the recent textbook by Klipp [52], while a more engineering-oriented treatment of modeling of biological circuits can be found in the text by Myers [69]. Two other books that are particularly noteworthy are Ptashne's book on the phage λ [73] and Madhani's book on yeast [58], both of which use well-studied model systems to describe a general set of mechanisms and principles that are present in many different types of organisms.

Several textbooks and research monographs provide excellent resources for modeling and analysis of biomolecular dynamics and regulation. J. D. Murray's two-volume text [67] on biological modeling is an excellent reference with many examples of biomolecular dynamics. The textbook by Phillips, Kondev and Theriot [72] provides a quantitative approach to understanding biological systems, including many of the concepts discussed in this chapter. Courey [17] gives a detailed

description of mechanisms transcriptional regulation.

The topics in dynamical systems and control theory that are briefly introduced here are covered in more detail in AM08 [1]. Other books that introduce tools for modeling and analysis of dynamical systems with applications in biology include J. D. Murray's text [67] and the recent text by and Ellner and Guckenheimer [22].

Synthetic biology is a rapidly evolving field that includes many different subareas of research, but few textbooks are currently available. In the specific area of biological circuit design that we focus on here, there are a number of good survey and review articles. The article by Baker *et al.* [7] provides a high level description of the basic approach and opportunities. Recent survey and review papers include Voigt [95] and Khalil and Collins [50].

Chapter 2
Dynamic Modeling of Core Processes

The goal of this chapter is to describe basic biological mechanisms in a way that can be represented by simple dynamical models. We begin the chapter with a discussion of the basic modeling formalisms that we will utilize to model biomolecular feedback systems. We then proceed to study a number of core processes within the cell, providing different model-based descriptions of the dynamics that will be used in later chapters to analyze and design biomolecular systems. The focus in this chapter and the next is on deterministic models using ordinary differential equations; Chapter 4 describes how to model the stochastic nature of biomolecular systems.

Prerequisites. Readers should have some basic familiarity with cell biology, at the level of the description in Section 1.2 (see also Appendix **??**), and a basic understanding of ordinary differential equations, at the level of Chapter 2 of AM08.

2.1 Modeling Techniques

In order to develop models for some of the core processes of the cell, we will need to build up a basic description of the biochemical reactions that take place, including production and degradation of proteins, regulation of transcription and translation, intracellular sensing, action and computation, and intercellular signaling. As in other disciplines, biomolecular systems can be modeled in a variety of different ways, at many different levels of resolution, as illustrated in Figure 2.1. The choice of which model to use depends on the questions that we want to answer, and good modeling takes practice, experience, and iteration. We must properly capture the aspects of the system that are important, reason about the appropriate temporal and spatial scales to be included, and take into account the types of simulation and analysis tools to be applied. Models that are to be used for analyzing existing systems should make testable predictions and provide insight into the underlying dynamics. Design models must additionally capture enough of the important behavior to allow decisions to be made regarding how to interconnect subsystems, choose parameters and design regulatory elements.

In this section we describe some of the basic modeling frameworks that we will build on throughout the rest of the text. We begin with brief descriptions of the relevant physics and chemistry of the system, and then quickly move to models that focus on capturing the behavior using reaction rate equations. In this chapter

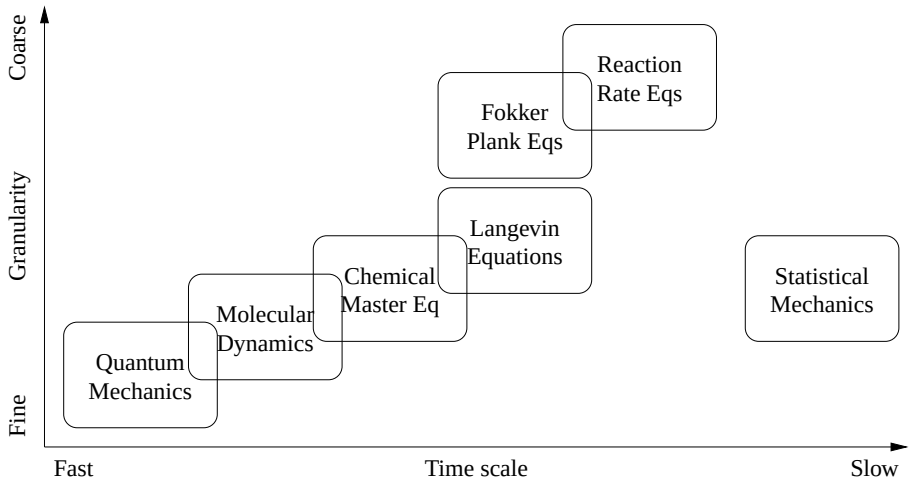


Figure 2.1: Different methods of modeling biomolecular systems.

our emphasis will be on dynamics with time scales measured in seconds to hours and mean behavior averaged across a large number of molecules. We touch only briefly on modeling in the case where stochastic behavior dominates and defer a more detailed treatment until Chapter 4.

Statistical mechanics and chemical kinetics

At the fine end of the modeling scale depicted in Figure 2.1, we can attempt to model the *molecular dynamics* of the cell, in which we attempt to model the individual proteins and other species and their interactions via molecular-scale forces and motions. At this scale, the individual interactions between protein domains, DNA and RNA are resolved, resulting in a highly detailed model of the dynamics of the cell.

For our purposes in this text, we will not require the use of such a detailed scale. Instead, we will start with the abstraction of molecules that interact with each other through stochastic events that are guided by the laws of thermodynamics. We begin with an equilibrium point of view, commonly referred to as *statistical mechanics*, and then briefly describe how to model the (statistical) dynamics of the system using chemical kinetics. We cover both of these points of view very briefly here, primarily as a stepping stone to deterministic models, and present a more detailed description in Chapter 4.

The underlying representation for both statistical mechanics and chemical kinetics is to identify the appropriate *microstates* of the system. A microstate corresponds to a given configuration of the components (species) in the system relative to each other and we must enumerate all possible configurations between the molecules that are being modeled.

Figure 2.2: Microstates for RNA polymerase. Each microstate of the system corresponds to the RNA polymerase being located at some position in the cell. If we discretize the possible locations on the DNA and in the cell, the microstates corresponds to all possible non-overlapping locations of the RNA polymerases. Figure from Phillips, Kondev and Theriot [72]; used with permission of Garland Science.

s an example, consider the distribution of RNA polymerase in the cell. It is known that most RNA polymerases are bound to the DNA in a cell, either as they produce RNA or as they diffuse along the DNA in search of a promoter site. Hence we can model the microstates of the RNA polymerase system as all possible locations of the RNA polymerase in the cell, with the vast majority of these corresponding to the RNA polymerase at some location on the DNA. This is illustrated in Figure 2.2. In statistical mechanics, we model the configuration of the cell by the probability that the system is in a given microstate. This probability can be calculated based on the energy levels of the different microstates. The laws of statistical mechanics state that if we have a set of microstates Q, then the steady state probability that the system is in a particular microstate q is given by

$$\mathbb{P}(q) = \frac{1}{Z} e^{-E_q/(k_B T)}, \tag{2.1}$$

where E_q is the energy associated with the microstate $q \in Q$, k_B is the Boltzmann constant, T is the temperature in degrees Kelvin, and Z is a normalizing factor, known as the *partition function*,

$$Z = \sum_{q \in Q} e^{-E_q/(k_B T)}.$$

(These formulas are described in more detail in Chapter 4.)

By keeping track of those microstates that correspond to a given system state (also called a *macrostate*), we can compute the overall probability that a given macrostate is reached. Thus, if we have a set of states $S \subset Q$ that correspond to a given macrostate, then the probability of being in the set S is given by

$$P(S) = \frac{1}{Z}\sum_{q\in S} e^{-E_q/(k_B T)} = \frac{\sum_{q\in S} e^{-E_q/(k_B T)}}{\sum_{q\in Q} e^{-E_q/(k_B T)}}. \tag{2.2}$$

This can be used, for example, to compute the probability that some RNA polymerase is bound to a given promoter, averaged over many independent samples, and from this we can reason about the rate of expression of the corresponding gene. More details and several examples will be illustrated in Chapter 4.

Statistical mechanics describes the steady state distribution of microstates, but does not tell us how the microstates evolve in time. To include the dynamics, we must consider the *chemical kinetics* of the system and model the probability that we transition from one microstate to another in a given period of time. Let q represent the microstate of the system, which we shall take as a vector of integers that represents the number of molecules of a specific types in given configurations or locations. Assume we have a set of M reactions R_j, $j = 1,\ldots,M$, with ξ_j representing the change in state q associated with reaction R_j. We describe the kinetics of the system by making use of the *propensity function* $a_j(q,t)$ associated with reaction R_j, which captures the instantaneous probability that at time t a system will transition between state q and state $q+\xi_j$.

More specifically, the propensity function is defined such that

$$a_j(q,t)dt \quad = \quad \text{Probability that reaction } R_j \text{ will occur between time } t \text{ and time } t+dt \text{ given that the microstate is } q.$$

We will give more detail in Chapter 4 regarding the validity of this functional form, but for now we simply assume that such a function can be defined for our system.

Using the propensity function, we can keep track of the probability distribution for the state by looking at all possible transitions into and out of the current state. Specifically, given $P(q,t)$, the probability of being in state q at time t, we can compute the time derivative $dP(q,t)/dt$ as

$$\frac{dP}{dt}(q,t) = \sum_{j=1}^{M}\Big(a_j(q-\xi_j)P(q-\xi_j,t) - a_j(q)P(q,t)\Big). \tag{2.3}$$

This equation (and its many variants) is called the *chemical master equation* (CME). The first sum on the right hand side represents the transitions into the state q from some other state $q-\xi_j$ and the second sum represents that transitions out of the state q. The variable ξ_j in the sum ranges over all possible reactions.

Clearly the dynamics of the distribution $P(q,t)$ depend on the form of the propensity functions $a_j(q)$. Consider a simple reaction of the form

$$\mathrm{A}+\mathrm{B} \rightleftharpoons \mathrm{AB} \quad \equiv \quad \begin{aligned} R_\mathrm{f} &: \mathrm{A}+\mathrm{B} \longrightarrow \mathrm{AB} \\ R_\mathrm{r} &: \mathrm{AB} \longrightarrow \mathrm{A}+\mathrm{B}. \end{aligned} \tag{2.4}$$

We assume that the reaction takes place in a well-stirred volume Ω and let the configurations q be represented by the number of each species that is present. The forward reaction R_f is a bimolecular reaction and we will see in Chapter 4 that it has a propensity function

$$a_\mathrm{f}(q) = \frac{k_\mathrm{f}}{\Omega} n_\mathrm{A} n_\mathrm{B},$$

where k_f is a parameter that depends on the forward reaction, and n_A and n_B are the number of molecules of each species. The reverse reaction R_r is a unimolecular reaction and we will see that it has a propensity function

$$a_\mathrm{r}(q) = k_\mathrm{r} n_\mathrm{AB},$$

where k_r is a parameter that depends on the reverse reaction and n_AB is the number of molecules of AB that are present.

If we now let $q = (n_\mathrm{A}, n_\mathrm{B}, n_\mathrm{AB})$ represent the microstate of the system, then we can write the chemical master equation as

$$\frac{dP}{dt}(n_\mathrm{A}, n_\mathrm{B}, n_\mathrm{AB}) = k_\mathrm{r} n_\mathrm{AB} P(n_\mathrm{A}-1, n_\mathrm{B}-1, n_\mathrm{AB}+1) - k_\mathrm{f} n_\mathrm{A} n_\mathrm{B} P(n_\mathrm{A}, n_\mathrm{B}, n_\mathrm{AB}).$$

The first term on the right hand side represents the transitions into the microstate $q = (n_\mathrm{A}, n_\mathrm{B}, n_\mathrm{AB})$ and the second term represents the transitions out of that state.

The number of differential equations depends on the number of molecules of A, B and AB that are present. For example, if we start with 1 molecules of A, 1 molecule of B, and 3 molecules of AB, then the possible states and dynamics are

$$\begin{aligned} q_0 &= (1,0,4) & dP_0/dt &= 3k_\mathrm{r} P_1 \\ q_1 &= (2,1,3) & dP_1/dt &= 4k_\mathrm{r} P_0 - 2(k_\mathrm{f}/\Omega) P_1 \\ q_2 &= (3,2,2) & dP_2/dt &= 3k_\mathrm{r} P_1 - 6(k_\mathrm{f}/\Omega) P_2 \\ q_3 &= (4,3,1) & dP_3/dt &= 2k_\mathrm{r} P_2 - 12(k_\mathrm{f}/\Omega) P_3 \\ q_4 &= (5,4,0) & dP_4/dt &= 1k_\mathrm{r} P_3 - 20(k_\mathrm{f}/\Omega) P_4, \end{aligned}$$

where $P_i = P(q_i, t)$. Note that the states of the chemical master equation are the probabilities that we are in a specific microstate, and the chemical master equation is a *linear* differential equation (we see from equation (2.3) that this is true in general).

The primary difference between the statistical mechanics description given by equation (2.1) and the chemical kinetics description in equation (2.3) is that the master

equation formulation describes how the probability of being in a given microstate evolves over time. Of course, if the propensity functions and energy levels are modeled properly, the steady state, average probabilities of being in a given microstate should be the same for both formulations.

Reaction rate equations

Although very general in form, the chemical master equation suffers from being a very high dimensional representation of the dynamics of the system. We shall see in Chapter 4 how to implement simulations that obey the master equation, but in many instances we will not need this level of detail in our modeling. In particular, there are many situations in which the number of molecules of a given species is such that we can reason about the behavior of a chemically reacting system by keeping track of the *concentration* of each species as a real number. This is of course an approximation, but if the number of molecules is sufficiently large, then the approximation will generally be valid and our models can be dramatically simplified.

To go from the chemical master equation to a simplified form of the dynamics, we begin by making a number of assumptions. First, we assume that we can represent the state of a given species by its concentration n_A/Ω, where n_A is the number of molecules of A in a given volume Ω. We also treat this concentration as a real number, ignoring the fact that the real concentration is quantized. Finally, we assume that our reactions take place in a well-stirred volume, so that the rate of interactions between two species is solely determined by the concentrations of the species.

Before proceeding, we should recall that in many (and perhaps most) situations inside of cells, these assumptions are *not* particularly good ones. Biomolecular systems often have very small molecular counts and are anything but well mixed. Hence, we should not expect that models based on these assumptions should perform well at all. However, experience indicates that in many cases the basic form of the equations provides a good model for the underlying dynamics and hence we often find it convenient to proceed in this manner.

Putting aside our potential concerns, we can now proceed to write the dynamics of a system consisting of a set of species S_i, $i = 1, \ldots, n$ undergoing a set of reactions R_j, $j = 1, \ldots, m$. We write $x_i = [S_i] = n_{S_i}/\Omega$ for the concentration of species i (viewed as a real number). Because we are interested in the case where the number of molecules is large, we no longer attempt to keep track of every possible configuration, but rather simply assume that the state of the system at any given time is given by the concentrations x_i. Hence the state space for our system is given by $x \in \mathbb{R}^n$ and we seek to write our dynamics in the form of a differential equation

$$\frac{dx}{dt} = f(x, \theta),$$

where $f : \mathbb{R}^n \to \mathbb{R}^n$ describes the rate of change of the concentrations as a function of the instantaneous concentrations and θ represents the parameters that govern the dynamic behavior.

To illustrate the general form of the dynamics, we consider again the case of a basic bimolecular reaction

$$\mathrm{A} + \mathrm{B} \rightleftharpoons \mathrm{AB}.$$

Each time the forward reaction occurs, we decrease the number of molecules of A and B by 1 and increase the number of molecules of AB (a separate species) by 1. Specifically, here AB denotes the complex formed by species A bound to species B. Similarly, each time the reverse reaction occurs, we decrease the number of molecules of AB by one and increase the number of molecules of A and B.

Using our discussion of the chemical master equation, we know that the likelihood that the forward reaction occurs in a given interval dt is given by $a_\mathrm{f}(q)dt = (k_\mathrm{f}/\Omega)n_\mathrm{A}n_\mathrm{B}dt$ and the reverse reaction has likelihood $a_\mathrm{r}(q) = k_\mathrm{r}n_\mathrm{AB}$. It follows that the concentration of the complex AB satisfies

$$\begin{aligned}
[\mathrm{AB}](t+dt) - [\mathrm{AB}](t) &= \mathbb{E}(n_\mathrm{AB}(t+dt)/\Omega - n_\mathrm{AB}(t)/\Omega) \\
&= \Big(a_\mathrm{f}(q - \xi_\mathrm{f}, t) - a_\mathrm{r}(q)\Big)/\Omega \cdot dt \\
&= \Big(k_\mathrm{f}n_\mathrm{A}n_\mathrm{B}/\Omega^2 - k_\mathrm{r}n_\mathrm{AB}/\Omega\Big)dt \\
&= \Big(k_\mathrm{f}[\mathrm{A}][\mathrm{B}] - k_\mathrm{r}[\mathrm{AB}]\Big)dt,
\end{aligned}$$

in which $\mathbb{E}(x)$ denotes the expected value of x. Taking the limit as dt approaches zero (but remains large enough that we can still average across multiple reactions, as described in more detail in Chapter 4), we obtain

$$\frac{d}{dt}[\mathrm{AB}] = k_\mathrm{f}[\mathrm{A}][\mathrm{B}] - k_\mathrm{r}[\mathrm{AB}].$$

In a similar fashion we can write equations to describe the dynamics of A and B and the entire system of equations is given by

$$\begin{aligned}
\frac{d}{dt}[\mathrm{A}] &= k_\mathrm{r}[\mathrm{AB}] - k_\mathrm{f}[\mathrm{A}][\mathrm{B}] \\
\frac{d}{dt}[\mathrm{B}] &= k_\mathrm{r}[\mathrm{AB}] - k_\mathrm{f}[\mathrm{A}][\mathrm{B}] \\
\frac{d}{dt}[\mathrm{AB}] &= k_\mathrm{f}[\mathrm{A}][\mathrm{B}] - k_\mathrm{r}[\mathrm{AB}]
\end{aligned}
\qquad \text{or} \qquad
\begin{aligned}
\frac{dA}{dt} &= k_\mathrm{r}C - k_\mathrm{f}A \cdot B \\
\frac{dB}{dt} &= k_\mathrm{r}C - k_\mathrm{f}A \cdot B \\
\frac{dC}{dt} &= k_\mathrm{f}A \cdot B - k_\mathrm{r}C,
\end{aligned}$$

where $C = [\mathrm{AB}]$, $A = [\mathrm{A}]$, and $B = [\mathrm{B}]$. These equations are known as the *mass action kinetics* or the *reaction rate equations* for the system. The parameters k_f and k_r are called the *rate constants* and they match the parameters that were used in the underlying propensity functions.

Note that the same rate constants appear in each term, since the rate of production of AB must match the rate of depletion of A and B and vice versa. We adopt the standard notation for chemical reactions with specified rates and write the individual reactions as

$$\mathrm{A} + \mathrm{B} \xrightarrow{k_\mathrm{f}} \mathrm{AB}, \qquad \mathrm{AB} \xrightarrow{k_\mathrm{r}} \mathrm{A} + \mathrm{B},$$

where k_f and k_r are the reaction rates. For bidirectional reactions we can also write

$$\mathrm{A} + \mathrm{B} \underset{k_\mathrm{r}}{\overset{k_\mathrm{f}}{\rightleftharpoons}} \mathrm{AB}.$$

It is easy to generalize these dynamics to more complex reactions. For example, if we have a reversible reaction of the form

$$\mathrm{A} + 2\,\mathrm{B} \underset{k_\mathrm{r}}{\overset{k_\mathrm{f}}{\rightleftharpoons}} 2\,\mathrm{C} + \mathrm{D},$$

where A, B, C and D are appropriate species and complexes, then the dynamics for the species concentrations can be written as

$$\begin{aligned} \frac{d}{dt}A &= k_\mathrm{r}C^2 \cdot D - k_\mathrm{f}A \cdot B^2, & \frac{d}{dt}C &= 2k_\mathrm{f}A \cdot B^2 - 2k_\mathrm{r}C^2 \cdot D, \\ \frac{d}{dt}B &= 2k_\mathrm{r}C^2 \cdot D - 2k_\mathrm{f}A \cdot B^2, & \frac{d}{dt}D &= k_\mathrm{f}A \cdot B^2 - k_\mathrm{r}C^2 \cdot D. \end{aligned} \tag{2.5}$$

Rearranging this equation, we can write the dynamics as

$$\frac{d}{dt}\begin{pmatrix} A \\ B \\ C \\ D \end{pmatrix} = \begin{pmatrix} -1 & 1 \\ -2 & 2 \\ 2 & -2 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} k_\mathrm{f}A \cdot B^2 \\ k_\mathrm{r}C^2 \cdot D \end{pmatrix}. \tag{2.6}$$

We see that in this decomposition, the first term on the right hand side is a matrix of integers reflecting the stoichiometry of the reactions and the second term is a vector of rates of the individual reactions.

More generally, given a chemical reaction consisting of a set of species S_i, $i = 1, \ldots, n$ and a set of reactions R_j, $j = 1, \ldots, m$, we can write the mass action kinetics in the form

$$\frac{dx}{dt} = Nv(x),$$

where $N \in \mathbb{R}^{n \times m}$ is the *stoichiometry matrix* for the system and $v(x) \in \mathbb{R}^m$ is the *reaction flux vector*. Each row of $v(x)$ corresponds to the rate at which a given reaction occurs and the corresponding column of the stoichiometry matrix corresponds to the changes in concentration of the relevant species. For example, for the

system in equation (2.6) we have

$$x = (A, B, C, D), \qquad N = \begin{pmatrix} -1 & 1 \\ -2 & 2 \\ 2 & -2 \\ 1 & -1 \end{pmatrix}, \qquad v(x) = \begin{pmatrix} k_f A \cdot B^2 \\ k_r C^2 \cdot D \end{pmatrix}.$$

As we shall see in the next chapter, the structured form of this equation will allow us to explore some of the properties of the dynamics of chemically reacting systems.

Sometimes, the following notation will be used to denote birth and death of species

$$\emptyset \xrightarrow{k_f} \mathrm{A}, \ \mathrm{A} \xrightarrow{k_r} \emptyset.$$

We attach to the first reaction the differential equation

$$\frac{dA}{dt} = k_f,$$

and to the second reaction we attach the differential equation

$$\frac{dA}{dt} = -k_r, A.$$

From a physical point of view, these reactions simplify the representation of more complex processes, such as production of proteins or degradation of proteins due to proteases.

Example 2.1 (Covalent modification of a protein)**.** Consider the set of reactions involved in the phosphorylation of a protein by a kinase, as shown in Figure 2.23. Let S represent the substrate, K represent the kinase and S^* represent the phosphorylated (activated) substrate. The sets of reactions illustrated in Figure 2.23 are

R_1: K + ATP $\longrightarrow$ K:ATP
R_2: K:ATP $\longrightarrow$ K + ATP
R_3: S + K:ATP $\longrightarrow$ S:K:ATP
R_4: S:K:ATP $\longrightarrow$ S + K:ATP
R_5: S:K:ATP $\longrightarrow$ S^*:K:ADP
R_6: S^*:K:ADP $\longrightarrow$ S^* + K:ADP
R_7: K:ADP $\longrightarrow$ K + ADP
R_8: K + ADP $\longrightarrow$ K:ADP.

We now write the kinetics for each reaction:

$v_1 = k_1$[K][ATP], $v_5 = k_5$[S:K:ATP],
$v_2 = k_2$[K:ATP], $v_6 = k_6$[S^*:K:ADP],
$v_3 = k_3$[S][K:ATP], $v_7 = k_7$[K:ADP],
$v_4 = k_4$[S:K:ATP], $v_8 = k_8$[K][ADP].

We treat [ATP] as a constant (regulated by the cell) and hence do not directly track its concentration. (If desired, we could similarly ignore the concentration of

ADP since we have chosen not to include the many additional reactions in which it participates.)

The kinetics for each species are thus given by

$$\begin{aligned}
\frac{d}{dt}[\mathrm{K}] &= -v_1 + v_2 + v_7 - v_8 & \frac{d}{dt}[\mathrm{K{:}ATP}] &= v_1 - v_2 - v_3 + v_4 \\
\frac{d}{dt}[\mathrm{S}] &= -v_3 + v_4 & \frac{d}{dt}[\mathrm{S{:}K{:}ATP}] &= v_3 - v_4 - v_5 \\
\frac{d}{dt}[\mathrm{S}^*] &= v_6 & \frac{d}{dt}[\mathrm{S}^*\mathrm{{:}K{:}ADP}] &= v_5 - v_6 \\
\frac{d}{dt}[\mathrm{ADP}] &= v_7 - v_8 & \frac{d}{dt}[\mathrm{K{:}ADP}] &= v_6 - v_7 + v_8.
\end{aligned}$$

Collecting these equations together and writing the state as a vector, we obtain

$$\underbrace{\frac{d}{dt}\begin{pmatrix} [\mathrm{K}] \\ [\mathrm{K{:}ATP}] \\ [\mathrm{S}] \\ [\mathrm{S{:}K{:}ATP}] \\ [\mathrm{S}^*] \\ [\mathrm{S}^*\mathrm{{:}K{:}ADP}] \\ [\mathrm{ADP}] \\ [\mathrm{K{:}ADP}] \end{pmatrix}}_{x} = \underbrace{\begin{pmatrix} -1 & 1 & 0 & 0 & 0 & 0 & 1 & -1 \\ 1 & -1 & 1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 & 0 & 1 & -1 & 1 \end{pmatrix}}_{N} \underbrace{\begin{pmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \\ v_5 \\ v_6 \\ v_7 \\ v_8 \end{pmatrix}}_{v(x)},$$

which is in standard stoichiometric form. ∇

Reduced order mechanisms

In this section, we look at the dynamics of some common reactions that occur in biomolecular systems. Under some assumptions on the relative rates of reactions and concentrations of species, it is possible to derive reduced order expressions for the dynamics of the system. We focus here on an informal derivation of the relevant results, but return to these examples in the next chapter to illustrate that the same results can be derived using a more formal and rigorous approach.

Simple binding reaction. Consider the reaction in which two species A and B bind reversibly to form a complex C=AB:

$$\mathrm{A} + \mathrm{B} \underset{d}{\overset{a}{\rightleftharpoons}} \mathrm{C}, \tag{2.7}$$

where a is the association rate constant and b is the dissociation rate constant. Assume that B is a species that is controlled by other reactions in the cell and that the total concentration of A is conserved, so that $A + C = [\mathrm{A}] + [\mathrm{AB}] = A_{\mathrm{tot}}$. If the

dynamics of this reaction are fast compared to other reactions in the cell, then the amount of A and C present can be computed as a (steady state) function of B.

To compute how A and C depend on the concentration of B at the steady state, we must solve for the equilibrium concentrations of A and C. The rate equation for C is given by

$$\frac{dC}{dt} = aB\cdot(A_{\text{tot}} - C) - dC.$$

By setting $dC/dt = 0$ and letting $K_{\text{d}} := d/a$, we obtain the expressions

$$C = \frac{A_{\text{tot}}(B/K_{\text{d}})}{1+(B/K_{\text{d}})}, \qquad A = \frac{A_{\text{tot}}}{1+(B/K_{\text{d}})}.$$

The constant K_{d} is called the *dissociation constant* of the reaction. Its inverse measures the affinity of A binding to B. The steady state value of C increases with B while the steady state value of A decreases with B as more of A is found in the complex C.

Note that when $B \approx K_{\text{d}}$, A and C have roughly equal concentration. Thus the higher the value of K_{d}, the more B is required for A to form the complex C. K_{d} has the units of concentration and it can be interpreted as the concentration of B at which half of the total number of molecules of A are associated with B. Therefore a high K_{d} represents a weak affinity between A and B, while a low K_{d} represents a strong affinity.

Cooperative binding reaction. Assume now that B binds to A only after dimerization, that is, only after binding another molecule of B. Then, we have that reactions (2.7) become

$$\text{B} + \text{B} \underset{k_2}{\overset{k_1}{\rightleftharpoons}} \text{B}_2, \qquad \text{B}_2 + \text{A} \underset{d}{\overset{a}{\rightleftharpoons}} \text{C}, \qquad A + C = A_{\text{tot}},$$

in which B_2 denotes the dimer of B. The corresponding ODE model is given by

$$\frac{dB_2}{dt} = 2k_1B^2 - 2k_2B_2 - aB_2\cdot(A_{\text{tot}} - C) + dC, \qquad \frac{dC}{dt} = aB_2\cdot(A_{\text{tot}} - C) - dC.$$

By setting $dB_2/dt = 0$, $dC/dt = 0$, and by defining $K_m := k_2/k_1$, we we obtain that

$$B_2 = B^2/K_m, \qquad C = \frac{A_{\text{tot}}(B_2/K_{\text{d}})}{1+(B_2/K_{\text{d}})}, \qquad A = \frac{A_{\text{tot}}}{1+(B_2/K_{\text{d}})},$$

so that

$$C = \frac{A_{\text{tot}}B^2/(K_mK_{\text{d}})}{1+B^2/(K_mK_{\text{d}})}, \qquad A = \frac{A_{\text{tot}}}{1+B^2/(K_mK_{\text{d}})}.$$

As an exercise, the reader can verify that if B binds to A only as a complex of n copies of B, that is,

$$\text{B} + \text{B} + \cdots + \text{B} \underset{k_2}{\overset{k_1}{\rightleftharpoons}} \text{B}_{\text{n}}, \qquad \text{B}_{\text{n}} + \text{A} \underset{d}{\overset{a}{\rightleftharpoons}} \text{C}, \qquad A + C = A_{\text{tot}},$$

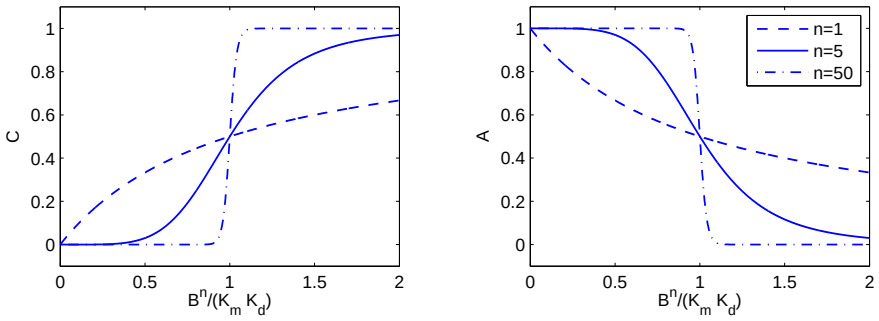


Figure 2.3: Steady state concentrations of the complex C and of A as functions of the concentration of B.

then we have that

$$C = \frac{A_{\text{tot}} B^n/(K_m K_{\text{d}})}{1 + B^n/(K_m K_{\text{d}})}, \qquad A = \frac{A_{\text{tot}}}{1 + B^n/(K_m K_{\text{d}})}.$$

In this case, one says that the binding of B to A is *cooperative* with cooperativity *n*. Figure 2.3 shows the above functions, which are often referred to as *Hill functions*.

Another type of cooperative binding is when a species R can bind A only after another species B as bound. In this case, the reactions are given by

$$\text{B} + \text{A} \underset{d}{\overset{a}{\rightleftharpoons}} \text{C}, \qquad \text{R} + \text{C} \underset{d'}{\overset{a'}{\rightleftharpoons}} \text{C}', \qquad A + C + C' = A_{\text{tot}}.$$

Proceeding as above by writing the ODE model and equating the time derivatives to zero to obtain the equilibrium, one obtains

$$C = \frac{1}{K_{\text{d}}} B(A_{\text{tot}} - C - C'), \qquad C' = \frac{1}{K'_{\text{d}} K_{\text{d}}} R(A_{\text{tot}} - C - C').$$

By solving this system of two equations for the unknowns C' and C, one obtains

$$C' = \frac{A_{\text{tot}}(B/K_{\text{d}})(R/K'_{\text{d}})}{1 + (B/K_{\text{d}}) + (B/K_{\text{d}})(R/K'_{\text{d}})}, \qquad C = \frac{A_{\text{tot}}(B/K_{\text{d}})}{1 + (B/K_{\text{d}}) + (B/K_{\text{d}})(R/K'_{\text{d}})}.$$

In the case in which B would first bind cooperatively with other copies of B with cooperativity n, the above expressions would modify to

$$C' = \frac{A_{\text{tot}}(B^n/K_{\text{d}}K_{\text{m}})(R/K'_{\text{d}})}{1 + (B^n/K_{\text{d}}K_{\text{m}})(R/K'_{\text{d}}) + (B^n/K_{\text{d}}K_{\text{m}})}, \qquad C = \frac{A_{\text{tot}}(B^n/K_{\text{d}}K_{\text{m}})}{1 + (B^n/K_{\text{d}}K_{\text{m}})(R/K'_{\text{d}}) + (B^n/K_{\text{d}}K_{\text{m}})}.$$

Competitive binding reaction. Finally, consider the case in which two species B_{a} and B_{r} both bind to A competitively, that is, they cannot be bound to A at the same

time. Let $\mathrm{C_a}$ be the complex formed between $\mathrm{B_a}$ and A and let $\mathrm{C_r}$ be the complex formed between $\mathrm{B_r}$ and A. Then, we have the following reactions

$$\mathrm{B_a} + \mathrm{A} \underset{d}{\overset{a}{\rightleftharpoons}} \mathrm{C_a}, \qquad \mathrm{B_r} + \mathrm{A} \underset{d'}{\overset{a'}{\rightleftharpoons}} \mathrm{C_r}, \qquad A + C_a + C_r = A_{\mathrm{tot}},$$

for which we can write the dynamics as

$$\frac{dC_a}{dt} = aB_a \cdot (A_{\mathrm{tot}} - C_a - C_r) - dC_a, \qquad \frac{dC_r}{dt} = a'B_r \cdot (A_{\mathrm{tot}} - C_a - C_r) - d'C_r.$$

By setting the derivatives to zero, we obtain that

$$C_a(aB_a + d) = aB_a(A_{\mathrm{tot}} - C_r), \qquad C_r(a'B_r + d') = a'B_r(A_{\mathrm{tot}} - C_a),$$

so that

$$C_r = \frac{B_r(A_{\mathrm{tot}} - C_a)}{B_r + K'_{\mathrm{d}}}, \qquad C_a\left(B_a + K_{\mathrm{d}} - \frac{B_a B_r}{B_r + K'_{\mathrm{d}}}\right) = B_a\left(\frac{K'_{\mathrm{d}}}{B_r + K'_{\mathrm{d}}}\right)A_{\mathrm{tot}},$$

from which we finally obtain that

$$C_a = \frac{A_{\mathrm{tot}}(B_a/K_{\mathrm{d}})}{1 + (B_a/K_{\mathrm{d}}) + (B_r/K'_{\mathrm{d}})}, \qquad C_r = \frac{A_{\mathrm{tot}}(B_r/K'_{\mathrm{d}})}{1 + (B_a/K_{\mathrm{d}}) + (B_r/K'_{\mathrm{d}})}.$$

In this derivation, we have assumed that both $\mathrm{B_a}$ and $\mathrm{B_r}$ bind A as monomers. If they were binding as dimers, the reader should verify as an exercise (see Exercises) that they would appear in the final expressions with a power of two.

Note also that in this derivation we have assumed that the binding is competitive, that is, $\mathrm{B_a}$ and $\mathrm{B_r}$ cannot simultaneously bind to A. If they were binding simultaneously to A, we would have included another complex comprising $\mathrm{B_a}$, $\mathrm{B_r}$ and A. Denoting this new complex by C′, we would have added also the two additional reactions

$$\mathrm{C_a} + \mathrm{B_r} \underset{\bar{d}}{\overset{\bar{a}}{\rightleftharpoons}} \mathrm{C}', \qquad \mathrm{C_r} + \mathrm{B_a} \underset{\bar{d}'}{\overset{\bar{a}'}{\rightleftharpoons}} \mathrm{C}'$$

and we would have modified the conservation law for A to $A_{\mathrm{tot}} = A + C_a + C_r + C'$. The reader can verify as an exercise (see Exercises) that in this case a mixed term $B_r B_a$ would appear in the equilibrium expressions.

Enzymatic reaction. A general enzymatic reaction can be written as

$$\mathrm{E} + \mathrm{S} \underset{d}{\overset{a}{\rightleftharpoons}} \mathrm{C} \overset{k}{\rightarrow} \mathrm{E} + \mathrm{P},$$

in which E is an enzyme, S is the substrate to which the enzyme binds to form the complex C=ES, and P is the product resulting from the modification of the substrate S due to the binding with the enzyme E. The parameter *a* is referred to as

association rate constant, d as dissociation rate constant, and k as the catalytic rate constant. Enzymatic reactions are very common and we will see specific instances of them in the sequel, e.g., phosphorylation and dephosphorylation reactions. The corresponding ODE system is given by

$$\frac{dS}{dt} = -aE \cdot S + dC, \qquad \frac{dC}{dt} = aE \cdot S - (d+k)C,$$
$$\frac{dE}{dt} = -aE \cdot S + dC + kC, \qquad \frac{dP}{dt} = kC.$$

The total enzyme concentration is usually constant and denoted by E_{tot}, so that $E + C = E_{\text{tot}}$. Substituting in the above equations $E = E_{\text{tot}} - C$, we obtain

$$\frac{dE}{dt} = -a(E_{\text{tot}} - C) \cdot S + dC + kC, \qquad \frac{dC}{dt} = a(E_{\text{tot}} - C) \cdot S - (d+k)C,$$
$$\frac{dS}{dt} = -a(E_{\text{tot}} - C) \cdot S + dC, \qquad \frac{dP}{dt} = kC.$$

This system cannot be solved analytically, therefore assumptions have been used in order to reduce it to a simpler form. Michaelis and Menten assumed that the conversion of E and S to C and *vice versa* is much faster than the decomposition of C into E and P. Under this assumption and letting $S(0)$ be sufficiently large (see Example 3.13), C immediately reaches its steady state value (while P is still changing). This approximation is called the *quasi-steady state assumption* and the mathematical conditions on the parameters that justify it will be dealt with in Section 3.6. The steady state value of C is given by solving $a(E_{\text{tot}} - C)S - (d+k)C = 0$ for C, which gives

$$C = \frac{E_{\text{tot}} S}{S + K_m}, \quad \text{with} \quad K_m = \frac{d+k}{a},$$

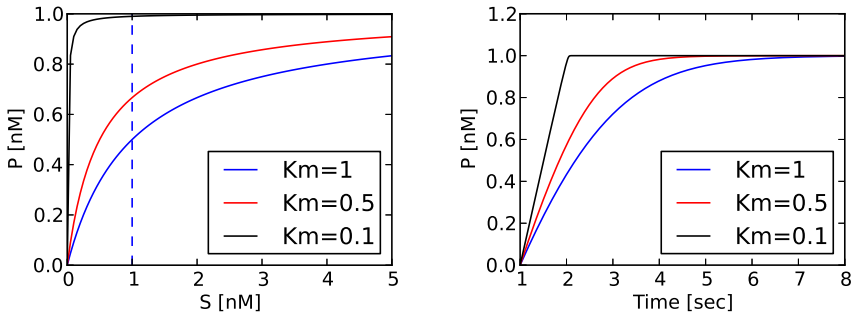
in which the constant K_m is called the *Michaelis-Menten constant*. Letting $V_{\max} = kE_{\text{tot}}$, the resulting kinetics

$$\frac{dP}{dt} = k\frac{E_{\text{tot}} S}{S + K_m} = V_{\max}\frac{S}{S + K_m}$$

is called *Michaelis-Menten kinetics*.

The constant V_{max} is called the maximal velocity (or maximal flux) of modification and it represents the maximal rate that can be obtained when the enzyme is completely saturated by the substrate. The value of K_m corresponds to the value of S that leads to a half-maximal value of the P production rate. When the enzyme complex can be neglected with respect to the total substrate amount S_{tot}, we have that $S_{\text{tot}} \approx S + P$, so that the above equation can be also re-written as

$$\frac{dP}{dt} = \frac{V_{max}(S_{tot} - P)}{(S_{tot} - P) + K_m}.$$



(a) Transfer curves (b) Simulations

Figure 2.4: Enzymatic reactions. (a) Transfer curve showing the production rate for P as a function of substrate concentration. (b) Time plots of product $P(t)$ for different values of the K_m. In the plots $S_{\text{tot}} = 1$ and $V_{max} = 1$. The black plot shows the behavior for a value of K_m much smaller that the total substrate amount S_{tot}. This corresponds to a constant product formation rate (at least before the substrate is almost all converted to product, that is, $S_{\text{tot}} - P \approx K_m$), which is referred to *zero-order kinetics*.

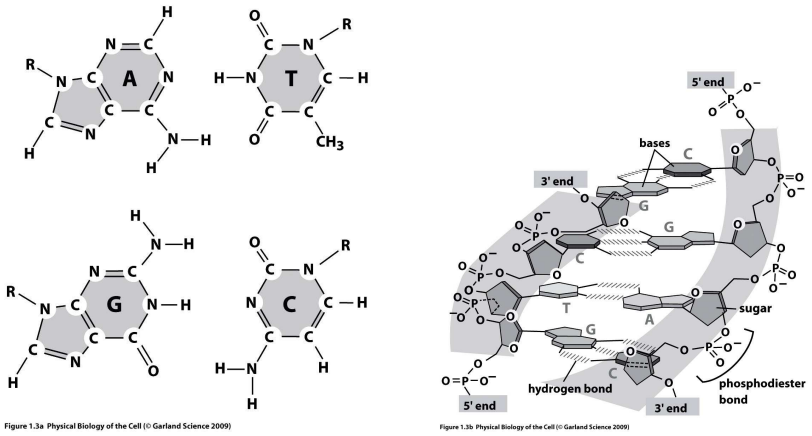
When $K_m \ll S_{\text{tot}}$ and the substrate has not yet been all converted to product, that is, $S_{\text{tot}} - P \gg K_m$, we have that the rate of product formation becomes approximately $dP/dt \approx V_{max}$, which is the maximal speed of reaction. Since this rate is constant and does not depend on the reactant concentrations, it is usually referred to *zero-order kinetics*. In this case, the system is said to operate in the zero-order regime (see Figure 2.4).

2.2 Transcription and Translation

In this section we consider the processes of transcription and translation, using the modeling techniques described in the previous section to capture the fundamental dynamic behavior. Models of transcription and translation can be done at a variety of levels of detail and which model to use depends on the questions that one wants to consider. We present several levels of modeling here, starting with a fairly detailed set of reactions and ending with highly simplified models that can be used when we are only interested in average production rate of proteins at relatively long time scales.

The central dogma: production of proteins

The genetic material inside a cell, encoded in its DNA, governs the response of a cell to various conditions. DNA is organized into collections of genes, with each gene encoding a corresponding protein that performs a set of functions in the cell. The activation and repression of genes are determined through a series of complex



(a) Base pairs (b) Double stranded

Figure 2.5: Molecular structure of DNA. (a) Individual bases (nucleotides) that make up DNA: adenine (A), cytocine (C), guanine (G) and thymine (T). (b) Double stranded DNA formed from individual nucleotides, with A binding to T and C binding to G. Each strand contains a 5' and 3' end, determined by the locations of the carbons where the next nucleotide binds. Figure from Phillips, Kondev and Theriot [72]; used with permission of Garland Science.

interactions that give rise to a remarkable set of circuits that perform the functions required for life, ranging from basic metabolism to locomotion to procreation. Genetic circuits that occur in nature are robust to external disturbances and can function in a variety of conditions. To understand how these processes occur (and some of the dynamics that govern their behavior), it will be useful to present a relatively detailed description of the underlying biochemistry involved in the production of proteins.

DNA is double stranded molecule with the "direction" of each strand specified by looking at the geometry of the sugars that make up its backbone (see Figure 2.5). The complementary strands of DNA are composed of a sequence of nucleotides that consist of a sugar molecule (deoxyribose) bound to one of 4 bases: adenine (A), cytocine (C), guanine (G) and thymine (T). The coding strand (by convention the top row of a DNA sequence when it is written in text form) is specified from the 5' end of the DNA to the 3' end of the DNA. (As described briefly in Appendix **??**, 5' and 3' refer to carbon locations on the deoxyribose backbone that are involved in linking together the nucleotides that make up DNA.) The DNA that encodes proteins consists of a promoter region, regulator regions (described in more detail below), a coding region and a termination region (see Figure 2.6). We informally refer to this entire sequence of DNA as a gene.

Expression of a gene begins with the *transcription* of DNA into mRNA by RNA polymerase, as illustrated in Figure 2.7. RNA polymerase enzymes are present in

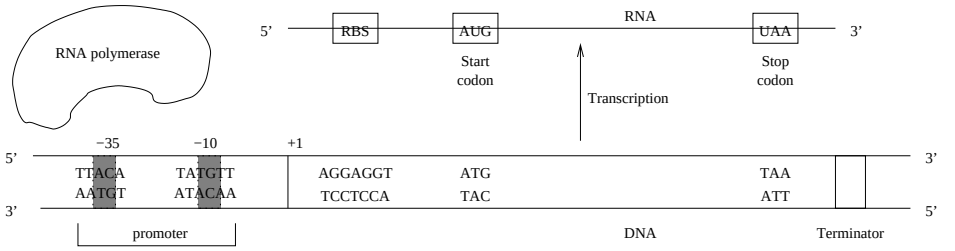


Figure 2.6: Geometric structure of DNA. The layout of the DNA is shown at the top. RNA polymerase binds to the promoter region of the DNA and transcribes the DNA starting at the +1 side and continuing to the termination site.

./coreproc/figures/Cou08_02_01.eps

Figure 2.7: Production of messenger RNA from DNA. RNA polymerase, along with other accessory factors, binds to the promoter region of the DNA and then "opens" the DNA to begin transcription (initiation). As RNA polymerase moves down the DNA, producing an RNA transcript (elongation), which is later translated into a protein. The process ends when the RNA polymerase reaches the terminator (termination). Reproduced from Courey [17]; permission pending.

the nucleus (for eukaryotes) or cytoplasm (for prokaryotes) and must localize and bind to the promoter region of the DNA template. Once bound, the RNA polymerase "opens" the double stranded DNA to expose the nucleotides that make up the sequence. This reversible reaction, called *isomerization*, is said to transform the RNA polymerase and DNA from a *closed complex* to an *open complex*. After the open complex is formed, RNA polymerase begins to travel down the DNA strand and constructs an mRNA sequence that matches the 5' to 3' sequence of the DNA to which it is bound. By convention, we number the first base pair that is transcribed as '+1' and the base pair prior to that (which is not transcribed) is labeled as '-1'. The promoter region is often shown with the -10 and -35 regions indicated, since these regions contain the nucleotide sequences to which the RNA polymerase enzyme binds (the locations vary in different cell types, but these two numbers are typically used).

The RNA strand that is produced by RNA polymerase is also a sequence of nucleotides with a sugar backbone. The sugar for RNA is ribose instead of deoxyribose and mRNA typically exists as a single stranded molecule. Another difference is that the base thymine (T) is replaced by uracil (U) in RNA sequences. RNA polymerase produces RNA one base pair at a time, as it moves from in the 5' to 3' direction along the DNA coding strand. RNA polymerase stops transcribing DNA when it reaches a *termination region* (or *terminator*) on the DNA. This termination region consists of a sequence that causes the RNA polymerase to unbind from the DNA. The sequence is not conserved across species and in many cells the termination sequence is sometimes "leaky", so that transcription will occasionally occur across the terminator.

Once the mRNA is produced, it must be translated into a protein. This process is slightly different in prokaryotes and eukaryotes. In prokaryotes, there is a region of the mRNA in which the ribosome (a molecular complex consisting of of both proteins and RNA) binds. This region, called the *ribosome binding site (RBS)*, has some variability between different cell species and between different genes in a given cell. The Shine-Delgarno sequence, AGGAGG, is the consensus sequence for the RBS. (A consensus sequence is a pattern of nucleotides that implements a given function across multiple organisms; it is not exactly conserved, so some variations in the sequence will be present from one organism to another.)

In eukaryotes, the RNA must undergo several additional steps before it is translated. The RNA sequence that has been created by RNA polymerase consists of *introns* that must be spliced out of the RNA (by a molecular complex called the spliceosome), leaving only the *exons*, which contain the coding sequence for the protein. The term *pre-mRNA* is often used to distinguish between the raw transcript and the spliced mRNA sequence, which is called *mature mRNA*. In addition to splicing, the mRNA is also modified to contain a *poly(A)* (polyadenine) *tail*, consisting of a long sequence of adenine (A) nucleotides on the 3' end of the mRNA. This processed sequence is then transported out of the nucleus into the cytoplasm,

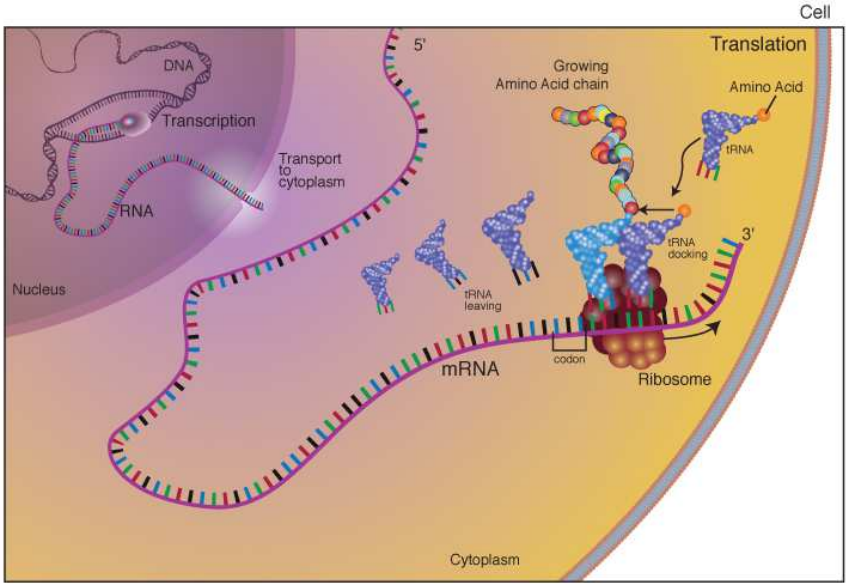


Figure 2.8: Translation is the process of translating the sequence of a messenger RNA (mRNA) molecule to a sequence of amino acids during protein synthesis. The genetic code describes the relationship between the sequence of base pairs in a gene and the corresponding amino acid sequence that it encodes. In the cell cytoplasm, the ribosome reads the sequence of the mRNA in groups of three bases to assemble the protein. Figure and caption courtesy the National Human Genome Research Institute.

where the ribosomes can bind to it.

Unlike prokaryotes, eukaryotes do not have a well defined ribosome binding sequence and hence the process of the binding of the ribosome to the mRNA is more complicated. The *Kozak sequence* A/GCCACCAUGG is the rough equivalent of the ribosome binding site, where the underlined AUG is the start codon (described below). However, mRNA lacking the Kozak sequence can also be translated.

Once the ribosome is bound to the mRNA, it begins the process of *translation*. Proteins consist of a sequence of amino acids, with each amino acid specified by a codon that is used by the ribosome in the process of translation. Each codon consists of three base pairs and corresponds to one of the 20 amino acids or a "stop" codon. The genetic code mapping between codons and amino acids is shown in Table **??**. The ribosome translates each codon into the corresponding amino acid using transfer RNA (tRNA) to integrate the appropriate amino acid (which binds to the tRNA) into the polypeptide chain, as shown in Figure 2.8. The start codon (AUG) specifies the location at which translation begins, as well as coding for the amino acid methionine (a modified form is used in prokaryotes). All subsequent codons are translated by the ribosome into the corresponding amino acid until it reaches one of the stop codons (typically UAA, UAG and UGA).

Table 2.1: Rates of core processes involved in the creation of proteins from DNA in *E. coli*.

Process	Characteristic rate	Source
mRNA transcription rate	24-29 bp/sec	BioNumbers [11]
Protein translation rate	12–21 aa/sec	BioNumbers [11]
Maturation time (fluorescent proteins)	6–60 min	BioNumbers [11]
mRNA half life	~ 100 sec	YM03 [99]
E. coli cell division time	20–40 min	BioNumbers [11]
Yeast cell division time	70–140 min	BioNumbers [11]
Protein half life	$\sim 5 \times 10^4$ sec	YM03 [99]
Protein diffusion along DNA	up to 10^4 bp/sec	PKT [72]

The sequence of amino acids produced by the ribosome is a polypeptide chain that folds on itself to form a protein. The process of folding is complicated and involves a variety of chemical interactions that are not completely understood. Additional post-translational processing of the protein can also occur at this stage, until a folded and functional protein is produced. It is this molecule that is able to bind to other species in the cell and perform the chemical reactions that underly the behavior of the organism. The *maturation time* of a protein is the time required for the polypeptide chain to fold into a functional protein.

Each of the processes involved in transcription, translation and folding of the protein takes time and affects the dynamics of the cell. Table 2.1 shows the rates of some of the key processes involved in the production of proteins. It is important to note that each of these steps is highly stochastic, with molecules binding together based on some propensity that depends on the binding energy but also the other molecules present in the cell. In addition, although we have described everything as a sequential process, each of the steps of transcription, translation and folding are happening simultaneously. In fact, there can be multiple RNA polymerases that are bound to the DNA, each producing a transcript. In prokaryotes, as soon as the ribosome binding site has been transcribed, the ribosome can bind and begin translation. It is also possible to have multiple ribosomes bound to a single piece of mRNA. Hence the overall process can be extremely stochastic and asynchronous.

Reaction models

The basic reactions that underly transcription include the diffusion of RNA polymerase from one part of the cell to the promoter region, binding of an RNA polymerase to the promoter, isomerization from the closed complex to the open complex, and finally the production of mRNA, one base pair at a time. To capture this set of reactions, we keep track of the various forms of RNA polymerase according to its location and state: $\mathrm{RNAP}^{\mathrm{c}}$ represents RNA polymerase in the cytoplasm, $\mathrm{RNAP}^{\mathrm{p}}$ represents RNA polymerase in the promoter region, and $\mathrm{RNAP}^{\mathrm{d}}$ is RNA

polymerase non-specifically bound to DNA. We must similarly keep track of the state of the DNA, to insure that multiple RNA polymerases do not bind to the same section of DNA. Thus we can write $\mathrm{DNA}^{\mathrm{p}}$ for the promoter region, $\mathrm{DNA}^{\mathrm{g},i}$ for the ith section of a gene g (whose length can depend on the desired resolution) and $\mathrm{DNA}^{\mathrm{t}}$ for the termination sequence. We write RNAP:DNA to represent RNA polymerase bound to DNA (assumed closed) and $\mathrm{RNAP{:}DNA}^{\mathrm{o}}$ to indicate the open complex. Finally, we must keep track of the mRNA that is produced by transcription: we write $\mathrm{mRNA}^{\mathrm{i}}$ to represent an mRNA strand of length i and assume that the length of the gene of interest is N.

Using these various states of the RNA polymerase and locations on the DNA, we can write a set of reactions modeling the basic elements of transcription as

$$
\begin{aligned}
\text{Binding to DNA:} \quad & \mathrm{RNAP}^{\mathrm{c}} \rightleftharpoons \mathrm{RNAP}^{\mathrm{d}} \\
\text{Diffusion along DNA:} \quad & \mathrm{RNAP}^{\mathrm{d}} \rightleftharpoons \mathrm{RNAP}^{\mathrm{p}} \\
\text{Binding to promoter:} \quad & \mathrm{RNAP}^{\mathrm{p}} + \mathrm{DNA}^{\mathrm{p}} \rightleftharpoons \mathrm{RNAP{:}DNA}^{\mathrm{p}} \\
\text{Isomerization:} \quad & \mathrm{RNAP{:}DNA}^{\mathrm{p}} \longrightarrow \mathrm{RNAP{:}DNA}^{\mathrm{o}} \\
\text{Start of transcription:} \quad & \mathrm{RNAP{:}DNA}^{\mathrm{o}} \longrightarrow \mathrm{RNAP{:}DNA}^{\mathrm{g},1} + \mathrm{DNA}^{\mathrm{p}} \\
\text{mRNA creation:} \quad & \mathrm{RNAP{:}DNA}^{\mathrm{g},1} \longrightarrow \mathrm{RNAP{:}DNA}^{\mathrm{g},2} + \mathrm{mRNA}_{\mathrm{k}}^{1} \\
\text{Elongation:} \quad & \mathrm{RNAP{:}DNA}^{\mathrm{g},\mathrm{i}+1} + \mathrm{mRNA}_{\mathrm{k}}^{\mathrm{i}} \\
& \qquad \longrightarrow \mathrm{RNAP{:}DNA}^{\mathrm{g},\mathrm{i}+2} + \mathrm{mRNA}_{\mathrm{k}}^{\mathrm{i}+1} \\
\text{Binding to terminator:} \quad & \mathrm{RNAP{:}DNA}^{\mathrm{g},\mathrm{N}} + \mathrm{mRNA}_{\mathrm{k}}^{\mathrm{N}-1} \\
& \qquad \longrightarrow \mathrm{RNAP{:}DNA}^{\mathrm{t}} + \mathrm{mRNA}_{\mathrm{k}}^{\mathrm{N}} \\
\text{Termination:} \quad & \mathrm{RNAP{:}DNA}^{\mathrm{t}} \longrightarrow \mathrm{RNAP}^{\mathrm{c}} \\
\text{Degradation:} \quad & \mathrm{mRNA}_{\mathrm{k}}^{\mathrm{N}} \longrightarrow \emptyset.
\end{aligned}
\tag{2.8}
$$

Not all these reactions occur on the same time scale. For example, the binding to promoter reaction is usually much faster than the isomerization reaction. Note that at the start of transcription we "release" the promoter region of the DNA, thus allowing a second RNA polymerase to bind to the promoter while the first RNA polymerase is still transcribing the gene. These reactions have been written for prokaryotes, but a similar set of reactions could be written for eukaryotes: the main differences would be that the RNA polymerase remains in the nucleus and the mRNA must be spliced and transported to the cytosol.

A similar set of reactions can be written to model the process of translation. Here we must keep track of the binding of the ribosome to the mRNA, translation of the mRNA sequence into a polypeptide chain, and folding of the polypeptide chain into a functional protein. Let $\mathrm{Ribo{:}mRNA}^{\mathrm{RBS}}$ indicate the ribosome bound to the ribosome binding site, $\mathrm{Ribo{:}mRNA}^{\mathrm{AA}i}$ the ribosome bound to the ith codon, $\mathrm{Ribo{:}mRNA}^{\mathrm{start}}$ and $\mathrm{Ribo{:}mRNA}^{\mathrm{stop}}$ for the start and stop codons, and PPC^{i} for a

polypeptide chain consisting of i amino acids. The reactions describing translation can then be written as

$$
\begin{aligned}
\text{Binding to RBS:} \quad & \text{Ribo} + \text{mRNA}_k^{\text{RBS}} \rightleftharpoons \text{Ribo:mRNA}_k^{\text{RBS}} \\
\text{Start of translation:} \quad & \text{Ribo:mRNA}_k^{\text{RBS}} \longrightarrow \text{Ribo:mRNA}_k^{\text{start}} + \text{mRNA}_k^{\text{RBS}} \\
\text{Polypeptide chain creation:} \quad & \text{Ribo:mRNA}_k^{\text{start}} \longrightarrow \text{Ribo:mRNA}_k^{\text{AA2}} + \text{PPC}^1 \\
\text{Elongation, } i = 1, \ldots, M\text{:} \quad & \text{Ribo:mRNA}_k^{\text{AA(i+1)}} + \text{PPC}^{\text{i}} \\
& \qquad \longrightarrow \text{Ribo:mRNA}_k^{\text{AA(i+2)}} + \text{PPC}^{\text{i+1}} \\
\text{Stop codon:} \quad & \text{Ribo:mRNA}_k^{\text{M}} + \text{PPC}^{\text{M-1}} \\
& \qquad \longrightarrow \text{Ribo:mRNA}_k^{\text{stop}} + \text{ppc}^{\text{M}} \\
\text{Release of mRNA:} \quad & \text{Ribo:mRNA}_k^{\text{stop}} \longrightarrow \text{Ribo} \\
\text{Folding:} \quad & \text{PPC}^{\text{M}} \longrightarrow \text{protein} \\
\text{Degradation:} \quad & \text{protein} \longrightarrow \emptyset.
\end{aligned}
$$

As in the case of transcription, we see that these reactions allow multiple ribosomes to translate the same piece of mRNA by freeing up the ribosome binding site (RBS) when translation begins.

As complex as these reactions are, they are still missing many important effects. For example, we have not accounted for the existence and effects of the 5' and 3' untranslated regions (UTRs) of a gene and we have also left out various error correction mechanisms in which ribosomes can step back and release an incorrect amino acid that has been incorporated into the polypeptide chain. We have also left out the many chemical species that must be present in order for a variety of the reactions to happen (NTPs for mRNA production, amino acids for protein production, etc). Incorporation of these effects requires additional reactions that track the many possible states of the molecular machinery that underlies transcription and translation.

When the details of the isomerization, start of transcription (translation), elongation, and terminaion are not relevant for the pheomenon to be studied, the transcription and translation reactions are lumped into much simpler reduced reactions. For the transcription process, these reduced reactions take the form:

$$
\begin{aligned}
&\text{RNAP:DNA}^{\text{p}} \xrightarrow{k_f} \text{mRNA} + \text{RNAP} + \text{DNA}^{\text{p}} \\
&\text{mRNA} \xrightarrow{\delta} \emptyset,
\end{aligned}
\tag{2.9}
$$

in which the first reaction lumps together isomerization, start of transcription, elongation, mRNA creation, and termination. Similarly, for the translation process, the

reduced reactions take the form:

$$\begin{aligned} &\text{Ribo} + \text{mRNA} \rightleftharpoons \text{Ribo:mRNA} \\ &\text{Ribo:mRNA} \longrightarrow \text{protein} + \text{mRNA} + \text{Ribo} \\ &\text{Ribo:mRNA} \longrightarrow \text{Ribo} \\ &\text{protein} \longrightarrow \emptyset, \end{aligned} \tag{2.10}$$

in which the second reaction lumps the start of translation, elongation, folding, and termination. The third reaction models the fact that mRNA can also be degraded when bound to ribosomes. The process of mRNA degradation occurs through RNAse enzymes binding to the ribosome binding site and cleaving the mRNA strand. It is known that ribosome binding site cannot be both bound to the ribosome and to the RNase [62]. However, the species Ribo:mRNA is a lumped species encompassing also configurations in which ribosomes are bound on the mRNA strand but not on the ribosome binding site. Hence, we also let this species be degraded by RNase.

Reaction rate equations

Given a set of reactions, the various stochastic processes that underly detailed models of transcription and translation can be specified using the stochastic modeling framework described briefly in the previous section. In particular, using either models of binding energy or measured rates, we can construct propensity functions for each of the many reactions that lead to production of proteins, including the motion of RNA polymerase and the ribosome along DNA and RNA. For many problems in which the detailed stochastic nature of the molecular dynamics of the cell are important, these models are the most relevant and they are covered in some detail in Chapter 4.

Alternatively, we can move to the reaction rate formalism and model the reactions using differential equations. To do so, we must compute the various reaction rates, which can be obtained from the propensity functions or measured experimentally. In moving to this formalism, we approximate the concentrations of various species as real numbers, which may not be accurate since some species exist at low molecular counts in the cell. Despite all of these approximations, in many situations the reaction rate equations are perfectly sufficient, particularly if we are interested in the average behavior of a large number of cells.

In some situations, an even simpler model of the transcription, translation and folding processes can be utilized. Let the "active" mRNA be the mRNA that is available for translation by the ribosome. We model its concentration through a simple time delay of length τ^m that accounts for the transcription of the ribosome binding site in prokaryotes or splicing and transport from the nucleus in eukaryotes. If we assume that RNA polymerase binds to DNA at some average rate (which includes both the binding and isomerization reactions) and that transcription takes

some fixed time (depending on the length of the gene), then the process of transcription can be described using the delay differential equation

$$\frac{dm_P}{dt} = \alpha - \mu m_P - \bar{\delta} m_P, \qquad m_P^*(t) = e^{-\mu\tau^m} m_P(t-\tau^m), \tag{2.11}$$

where m_P is the concentration of mRNA for protein P, m_P^* is the concentration of active mRNA, α is the rate of production of the mRNA for protein P, μ is the growth rate of the cell (which results in dilution of the concentration) and $\bar{\delta}$ is the rate of degradation of the mRNA. Since the dilution and degradation terms are of the same form, we will often combine these terms in the mRNA dynamics and use a single coefficient $\delta = \mu + \bar{\delta}$. The exponential factor accounts for dilution due to the change in volume of the cell, where μ is the cell growth rate. The constants α and δ capture the average rates of production and degradation, which in turn depend on the more detailed biochemical reactions that underlie transcription.

Once the active mRNA is produced, the process of translation can be described via a similar ordinary differential equation that describes the production of a functional protein:

$$\frac{dP}{dt} = \kappa m_P^* - \gamma P, \qquad P^f(t) = e^{-\mu\tau^f} P(t-\tau^f). \tag{2.12}$$

Here P represents the concentration of the polypeptide chain for the protein, P^f represents the concentration of functional protein (after folding). The parameters that govern the dynamics are κ, the rate of translation of mRNA; γ, the rate of degradation and dilution of P; and τ^f, the time delay associated with folding and other processes required to make the protein functional. The exponential term again accounts for dilution due to cell growth. The degradation and dilution term, parameterized by γ, captures both rate at which the polypeptide chain is degraded and the rate at which the concentration is diluted due to cell growth.

It will often be convenient to write the dynamics for transcription and translation in terms of the functional mRNA and functional protein. Differentiating the expression for m_P^*, we see that

$$\begin{aligned} \frac{dm_P^*(t)}{dt} &= e^{-\mu\tau^m} \dot{m}_P(t-\tau^m) \\ &= e^{-\mu\tau^m}(\alpha - \delta m_P(t-\tau^m)) = \bar{\alpha} - \delta m_P^*(t), \end{aligned} \tag{2.13}$$

where $\bar{\alpha} = e^{-\mu\tau^m}\alpha$. A similar expansion for the active protein dynamics yields

$$\frac{dP^f(t)}{dt} = \bar{\kappa} m_P^*(t-\tau^f) - \gamma P^f(t), \tag{2.14}$$

where $\bar{\kappa} = e^{-\mu\tau^f}\kappa$. We shall typically use equations (2.13) and (2.14) as our (reduced) description of protein folding, dropping the superscript f and overbars when there is no risk of confusion. Also, in the presence of different proteins, we will attach subscripts to the parameters to denote the protein they refer to.

In many situations the time delays described in the dynamics of protein production are small compared with the time scales at which the protein concentration changes (depending on the values of the other parameters in the system). In such cases, we can simplify our model of the dynamics of protein production even further and write

$$\frac{dm_P}{dt} = \alpha - \delta m_P, \qquad \frac{dP}{dt} = \kappa m_P - \gamma P. \tag{2.15}$$

Note that we here have dropped the superscripts $*$ and f since we are assuming that all mRNA is active and proteins are functional and dropped the overbar on α and κ since we are assuming the time delays are negligible.

Finally, the simplest model for protein production is one in which we only keep track of the basal rate of production of the protein, without including the mRNA dynamics. This essentially amounts to assuming the mRNA dynamics reach steady state quickly and replacing the first differential equation in equation (2.15) with its equilibrium value. This is often a good assumption as mRNA degration is usually about 100–1000 times faster than protein degradation (see Table 2.1). Thus we obtain

$$\frac{dP}{dt} = \beta - \gamma P, \qquad \beta := \kappa \frac{\alpha}{\delta}.$$

This model represents a simple first order, linear differential equation for the rate of production of a protein. In many cases this will be a sufficiently good approximate model, although we will see that in many cases it is too simple to capture the observed behavior of a biological circuit.

2.3 Transcriptional Regulation

The operation of a cell is governed in part by the selective expression of genes in the DNA of the organism, which control the various functions the cell is able to perform at any given time. Regulation of protein activity is a major component of the molecular activities in a cell. By turning genes on and off, and modulating their activity in more fine-grained ways, the cell controls the many metabolic pathways, responds to external stimuli, differentiates into different cell types as it divides, and maintains the internal state of the cell required to sustain life.

The regulation of gene expression and protein activity is accomplished through a variety of molecular mechanisms, as discussed in Section 1.2 and illustrated in Figure 2.9. At each stage of the processing from a gene to a protein, there are potential mechanisms for regulating the production processes. The remainder of this section will focus on transcriptional control and the next section on selected post-transcriptional control mechanisms. We will focus on prokaryotic mechanisms.

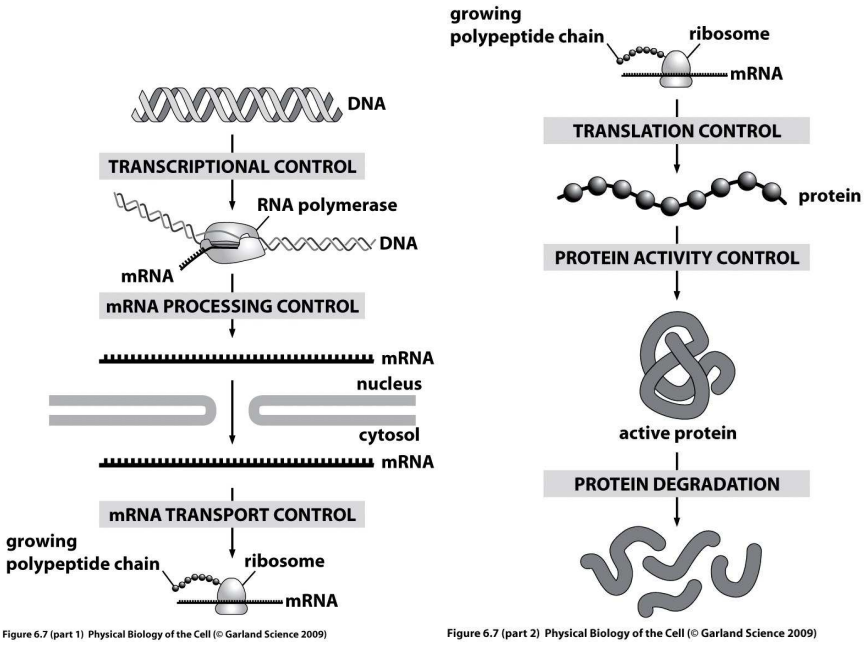


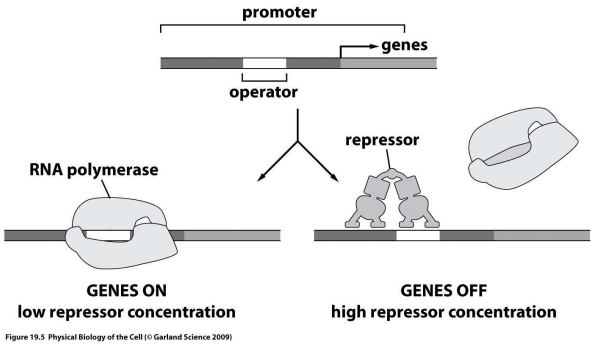
Figure 2.9: Regulation of proteins. Figure from Phillips, Kondev and Theriot [72]; used with permission of Garland Science.

Transcriptional regulation of protein production

There are a variety of mechanisms in the cell to regulate the production of proteins. These regulatory mechanisms can occur at various points in the overall process that produces the protein. Figure 2.9 shows some of the common points of regulation in the protein production process. We focus first on *transcriptional regulation*, which refers to regulatory mechanisms that control whether or not a gene is transcribed.

The simplest forms of transcriptional regulation are repression and activation, which are controlled through *transcription factors*. In the case of *repression*, the presence of a transcription factor (often a protein that binds near the promoter) turns off the transcription of the gene and this type of regulation is often called negative regulation or "down regulation". In the case of *activation* (or positive regulation), transcription is enhanced when an activator protein binds to the promoter site (facilitating binding of the RNA polymerase).

Represession. A common mechanism for repression is that a protein binds to a region of DNA near the promoter and blocks RNA polymerase from binding. The region of DNA to which the repressor protein binds is called an *operator region* (see Figure 2.10a). If the operator region overlaps the promoter, then the presence of a protein at the promoter can "block" the DNA at that location and transcription cannot initiate, as illustrated in Figure 2.10a. Repressor proteins often bind to



(a) Repression of gene expression

./coreproc/figures/PKT08_19_06.eps

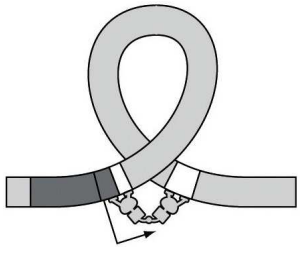
(b) Examples of repressors

Figure 2.10: Repression of gene expression. Figure from Phillips, Kondev and Theriot [72]; used with permission of Garland Science.

DNA as dimers or pairs of dimers (effectively tetramers). Figure 2.10b shows some examples of repressors bound to DNA.

A related mechanism for repression is *DNA looping*. In this setting, two repressor complexes (often dimers) bind in different locations on the DNA and then bind to each other. This can create a loop in the DNA and block the ability of RNA polymerase to bind to the promoter, thus inhibiting transcription. Figure 2.11 shows an example of this type of repression, in the *lac* operon. (An *operon* is a set of genes that is under control of a single promoter.)

Activation. The process of activation of a gene requires that an activator protein be

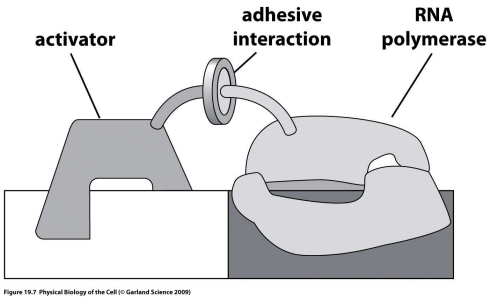


(a) DNA looping

./coreproc/figures/PKT08_08_19.eps

(b) *lac* repressor

Figure 2.11: Repression via DNA looping. Figure from Phillips, Kondev and Theriot [72]; used with permission of Garland Science.



(a) Activation mechanism

./coreproc/figures/PKT08_19_08.eps

(b) Examples of activators

Figure 2.12: Activation of gene expression. (a) Conceptual operation of an activator. The activator binds to DNA upstream of the gene and attracts RNA polymerase to the DNA strand. (b) Examples of activiators: catablite activator protein (CAP), p53 tumor supressor, zinc finger DNA binding domain and leucine zipper DAN binding domain. Figure from Phillips, Kondev and Theriot [72]; used with permission of Garland Science.

present in order for transcription to occur. In this case, the protein must work to either recruit or enable RNA polymerase to begin transcription.

The simplest form of activation involves a protein binding to the DNA near the promoter in such a way that the combination of the activator and the promoter sequence bind RNA polymerase. Figure 2.12 illustrates the basic concept. Like repressors, many activators have inducers, which can act in either a positive or negative fashion (see Figure 2.14b). For example, cyclic AMP (cAMP) acts as a positive inducer for CAP.

Another mechanism for activation of transcription, specific to prokaryotes, is the use of *sigma factors*. Sigma factors are part of a modular set of proteins that bind to RNA polymerase and form the molecular complex that performs transcription. Different sigma factors enable RNA polymerase to bind to different promoters, so the sigma factor acts as a type of activating signal for transcription. Table 2.2 lists some of the common sigma factors in bacteria. One of the uses of sigma factors is to produce certain proteins only under special conditions, such as when the cell undergoes *heat shock*. Another use is to control the timing of the expression of certain genes, as illustrated in Figure 2.13.

Inducers. A feature that is present in some types of transcription factors is the existence of an *inducer molecule* that combines with the protein to either activate or inactivate its function. A *positive inducer* is a molecule that must be present in order for repression or activation to occur. A *negative inducer* is one in which the presence of the inducer molecule blocks repression or activation, either by changing the

Table 2.2: Sigma factors in *E. coli* [2].

Sigma factor	Promoters recognized
σ^{70}	most genes
σ^{32}	genes associated with heat shock
σ^{28}	genes involved in stationary phase and stress response
σ^{28}	genes involved in motility and chemotaxis
σ^{24}	genes dealing with misfolded proteins in the periplasm

shape of the transcription factor protein or by blocking active sites on the protein that would normally bind to the DNA. Figure 2.14a summarizes the various possibilities. Common examples of repressor-inducer pairs include *lacI* and lactose (or IPTG), *tetR* and aTc, and tryptophan repressor and tryptophan. Lactose/IPTG and aTc are both negative inducers, so their presence causes the otherwise repressed gene to be expressed, while tryptophan is a positive inducer.

Combinatorial promoters. In addition to promoters that can take either a repressor or an activator as the sole input transcription factor, there are *combinatorial promoters* that can take both repressors and activators as input transcription factors. This allows genes to be switched on and off based on more complex conditions, represented by the concentrations of two or more activators or repressors.

Figure 2.15 shows one of the classic examples, a promoter for the *lac* system. In the *lac* system, the expression of genes for metabolizing lactose are under the control of a single (combinatorial) promoter. CAP, which is positively induced by cAMP, acts as an activator and LacI (also called "lac repressor"), which is negatively induced by lactose, acts as a repressor. In addition, the inducer cAMP is expressed only when glucose levels are low. The resulting behavior is that the proteins for metabolizing lactose are expressed only in conditions where there is no glucose (so CAP is active) *and* lactose is present.

More complicated combinatorial promoters can also be used to control tran-

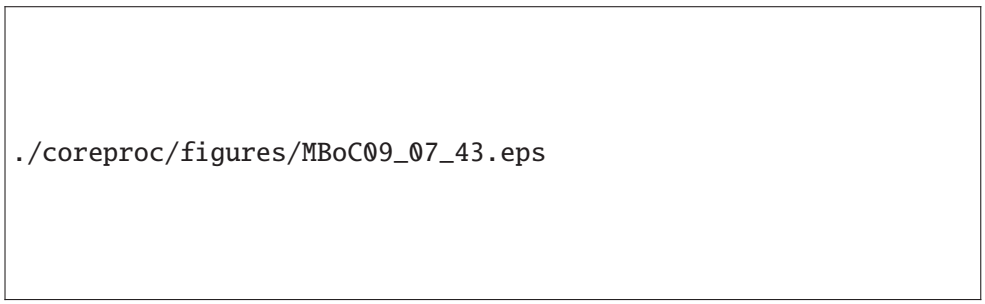


Figure 2.13: Use of sigma factors to controlling the timing of expression. Reproduced from Alberts et al. [2]; permission pending.

Figure 2.14: Effects of inducers. Reproduced from Alberts et al. [2]; permission pending.

scription in two different directions, an example that is found in some viruses.

Antitermination. A final method of activation in prokaryotes is the use of *antitermination*. The basic mechanism involves a protein that binds to DNA and deactivates a site that would normally serve as a termination site for RNA polymerase. Additional genes are located downstream from the termination site, but without a promoter region. Thus, in the presence of the anti-terminator protein, these genes are not expressed (or expressed with low probability). However, when the antitermination protein is present, the RNA polymerase maintains (or regains) its contact with the DNA and expression of the downstream genes is enhanced. In this way, antitermination allows downstream genes to be regulated by repressing "premature" termination. An example of an antitermination protein is the protein N in phage λ, which binds to a region of DNA labeled Nut (for N utilization), as shown in Figure 2.16 [35].

Reaction models

We can capture this set of molecular interactions by modifying the RNA polymerase binding reactions in equation (2.8). For a repressor (Rep), we simply have to add a reaction that represents the repressor bound to the promoter $\mathrm{DNA}^{\mathrm{p}}$:

$$\text{Repressor binding:} \quad \mathrm{DNA}^{\mathrm{p}} + \mathrm{Rep} \rightleftharpoons \mathrm{DNA{:}Rep}$$

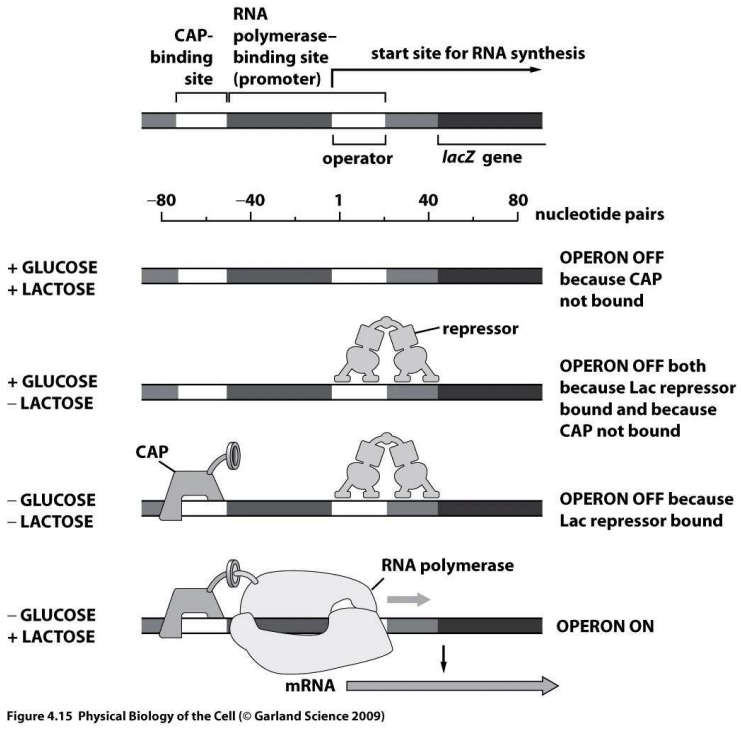


Figure 2.15: Combinatorial logic for the *lac* operator. Figure from Phillips, Kondev and Theriot [72]; used with permission of Garland Science.

This reaction acts to "sequester" the DNA promoter site so that it is no longer available for binding by RNA polymerase. The strength of the repressor is reflected in the reaction rate constants for the repressor binding reaction. Sometimes, the RNA polymerase can bind to the promoter even when the repressor is bound, usually with lower association rate constant. In this case, the repressor still allows some transcription even when bound to the promoter and the repressor is said to be "leaky".

The modifications for an activator (Act) are a bit more complicated, since we have to modify the reactions to require the presence of the activator before RNA

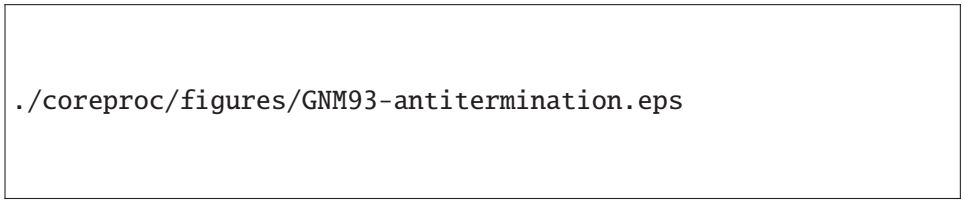


Figure 2.16: Antitermination. Reproduced from [35]; permission pending.

polymerase can bind the promoter. One possible mechanism, also known as the *recruitment model*, is given by

$$\begin{aligned}
\text{Activator binding:} &\quad \mathrm{DNA^p} + \mathrm{Act} \rightleftharpoons \mathrm{DNA^p{:}Act} \\
\text{RNAP binding w/ activator:} &\quad \mathrm{RNAP^p} + \mathrm{DNA^p{:}Act} \rightleftharpoons \mathrm{RNAP{:}DNA^p{:}Act} \\
\text{Isomerization:} &\quad \mathrm{RNAP{:}DNA^p{:}Act} \longrightarrow \mathrm{RNAP{:}DNA^o{:}Act} \\
\text{Start of transcription:} &\quad \mathrm{RNAP{:}DNA^o{:}Act} \longrightarrow \mathrm{RNAP{:}DNA^{g,1+}DNA^p{:}Act}.
\end{aligned} \tag{2.16}$$

In this model, RNA polymerase cannot bind to the promoter unless the activator is already bound to it. More generally, one can model both the enhanced binding of the RNA polymerase to the promoter in the presence of the activator, as well as the possibility of binding without an activator. This translates into the additional reaction $\mathrm{RNAP^p} + \mathrm{DNA^p} \rightleftharpoons \mathrm{RNAP{:}DNA^p}$. The relative reaction rates determine how strong the activator is and the "leakiness" of transcription in the absence of the activator. A different model of activation, called *allosteric*, is one in which the RNAP binding to DNA is not enhanced by the presence of the activator bound to the promoter, but the open complex (and hence start of transcription) formation can occur only (is enhanced) in the presence of the activator.

A simplified ordinary differential equation model can be obtained by accounting for the fact that transcription factors and RNAP bind to the DNA rapidly when compared to other reactions, such as isomerization, so that they can be well approximated by their quasi-steady state values. In this case, we can make use of the reduced order models described in Section 2.1. We can consider the competitive binding case to model a strong repressor that prevents RNAP from binding to the DNA. In the sequel, we remove the superscripts "p" and "d" from RNAP to simplify notation. The quasi-steady state concentration of the complex of DNA promoter bound to the repressor will have the expression

$$[\mathrm{DNA^p{:}Rep}] = \frac{[\mathrm{DNA}]([\mathrm{Rep}]/K_\mathrm{d})}{1 + [\mathrm{Rep}]/K_\mathrm{d} + [\mathrm{RNAP}]/K'_\mathrm{d}}$$

and the steady state amount of DNA promoter bound to the RNA polymerase will be given by

$$[\mathrm{RNAP{:}DNA^p}] = \frac{([\mathrm{RNAP}]/K'_\mathrm{d})[\mathrm{DNA}]}{1 + [\mathrm{RNAP}]/K'_\mathrm{d} + [\mathrm{Rep}]/K_\mathrm{d}},$$

in which K'_d is the dissociation constant of RNAP from the promoter while K_d is the dissociation constant of Rep from the promoter. The free promoter DNA with RNAP bound will allow transcription, while the complex $\mathrm{DNA^p}$:Rep will not allow transcription as it is not bound to RNAP. Using the lumped reactions (2.9), this can be modeled as

$$\frac{d[\mathrm{mRNA}]}{dt} = F([\mathrm{Rep}]) - \delta[\mathrm{mRNA}],$$

in which the production rate is given by

$$F([\text{Rep}]) = k_\text{f} \frac{[\text{DNA}]([\text{RNAP}]/K'_\text{d})}{1 + [\text{RNAP}]/K'_\text{d} + [\text{Rep}]/K_\text{d}}.$$

If the repressor binds to the promoter with cooperativity n, the above expression becomes (see Section 2.1)

$$F([\text{Rep}]) = k_\text{f} \frac{[\text{DNA}]([\text{RNAP}]/K'_\text{d})}{1 + [\text{RNAP}]/K'_\text{d} + [\text{Rep}]^n/(K_\text{d}K_\text{m})},$$

in which K_m is the dissociation constant of the reaction of n molecules of Rep binding together. The function F is usually denoted by the standard Hill function form

$$F([\text{Rep}]) = \frac{\alpha}{1 + ([\text{Rep}]/K)^n},$$

in which α and K are implicitly defined.

Finally, if the repressor allows RNAP to still bind to the promoter at a small rate (leaky repressor), the above expression can be modified to take the form (see Section 2.1)

$$F([\text{Rep}]) = \frac{\alpha}{1 + ([\text{Rep}]/K)^n} + \alpha_0, \tag{2.17}$$

in which α_0 is the basal expression level when the promoter is fully repressed, usually referred to as "leakiness".

To model the production rate of mRNA in the case in which an activator Act is required, we can consider first the case in which RNAP binds only when the activator is already bound to the promoter. To simplify the mathematical derivation, we re-write the reactions (2.16) involving the activator with the lumped transcription reaction (2.9) into the following:

$$\begin{aligned} &\text{DNA}^\text{p} + \text{Act} \rightleftharpoons \text{DNA}^\text{p}\text{:Act} \\ &\text{RNAP} + \text{DNA}^\text{p}\text{:Act} \rightleftharpoons \text{RNAP:DNA}^\text{p}\text{:Act} \\ &\text{RNAP:DNA}^\text{p}\text{:Act} \xrightarrow{k_\text{f}} \text{mRNA} + \text{RNAP} + \text{DNA}^\text{p}\text{:Act}. \end{aligned} \tag{2.18}$$

The first and second reactions fit the structure of the cooperative binding model illustrated in Section 2.1. Also, since the third reaction is much slower compared to the first two, the complex RNAP:DNA$^\text{p}$:Act concentration can be well approximated at its quasi-steady state. The expression of this quasi-steady state was given in Section 2.1 in correspondence to the cooperative binding model and takes the form:

$$[\text{RNAP:DNA}^\text{p}\text{:Act}] = \frac{([\text{RNAP}][\text{Act}])/(K_\text{d}K'_\text{d})[\text{DNA}]}{1 + ([\text{Act}]/K_\text{d})(1 + [\text{RNAP}]/K'_\text{d})},$$

in which K'_d is the dissociation constant of RNAP with the complex of DNA bound to Act and K_d is the dissociation constant of Act with DNA. When the activator Act

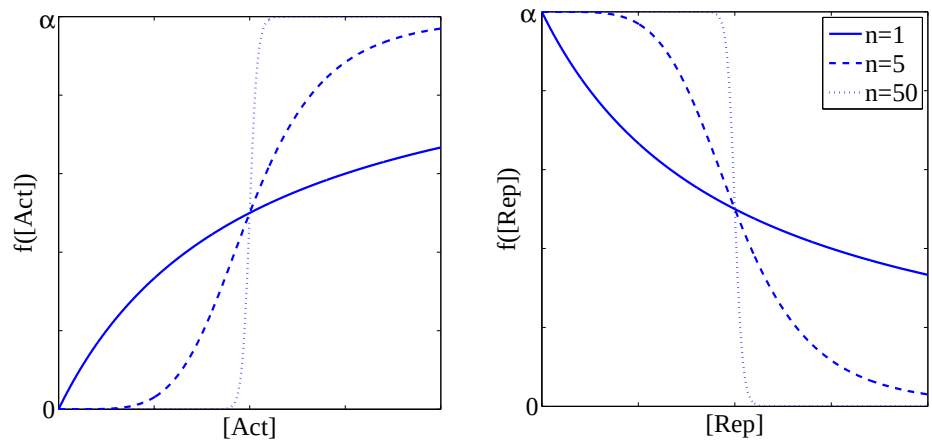


Figure 2.17: Hill function for an activator (left) and a repressor (right).

binds to the promoter with cooperativity n, the above expression will be modified to (see Section **??**):

$$[\text{RNAP:DNA}^{\text{p}}\text{:Act}] = \frac{([\text{RNAP}][\text{Act}]^n)/(K_\text{d}K'_\text{d}K_\text{m})[\text{DNA}]}{1+([\text{Act}]^n/K_\text{d}K_\text{m})(1+[\text{RNAP}]/K'_\text{d})},$$

in which K_m is the dissociation constant of the reaction of n molecules of Act binding together.

In order to write the differential equation for the mRNA concentration, we consider the third reaction in (2.18) along with the above quasi-steady state expressions of [RNAP:DNA$^{\text{p}}$:Act] to obtain

$$\frac{d\,[\text{mRNA}]}{dt} = F([\text{Act}]) - \delta[\text{mRNA}],$$

in which

$$F([\text{Act}]) = k_\text{f}\frac{([\text{RNAP}][\text{Act}]^n)/(K_\text{d}K'_\text{d}K_\text{m})[\text{DNA}]}{1+([\text{Act}]^n/K_\text{d}K_\text{m})(1+[\text{RNAP}]/K'_\text{d})} =: \frac{\alpha([\text{Act}]/K)^n}{1+([\text{Act}]/K)^n},$$

in which α and K are implicitly defined. The right-hand side expression is in the standard Hill function form. Figure 2.17 shows the shape of these Hill functions for both an activator and a repressor. If we assume that RNAP can still bind to DNA even when the activator is not bound, we have an additional basal expression rate α_0 so that the new form of the production rate is given by

$$F([\text{Act}]) = \frac{\alpha([\text{Act}]/K)^n}{1+([\text{Act}]/K)^n} + \alpha_0.$$

As indicated earlier, many activators and repressors operate in the presence of inducers. To incorporate these dynamics in our description, we simply have to add the reactions that correspond to the interaction of the inducer with the relevant protein. For a negative inducer, we can simply add a reaction in which the inducer binds the regulator protein and effectively sequesters it so that it cannot interact with the DNA. For example, a negative inducer operating on a repressor could be modeled by adding the reaction

$$\text{Rep} + \text{Ind} \rightleftharpoons \text{Rep:Ind}.$$

Since the above reactions are very fast compared to transcription, they can be assumed at the quasi-steady state. Hence, the free amount of repressor that can still bind to the promoter can be calculated by writing the ODE model corresponding to the above reactions and by setting the time derivatives to zero. This yields

$$[\text{Rep}] = \frac{[\text{Rep}]_{\text{tot}}}{1 + [\text{Ind}]/K_{\text{d}}},$$

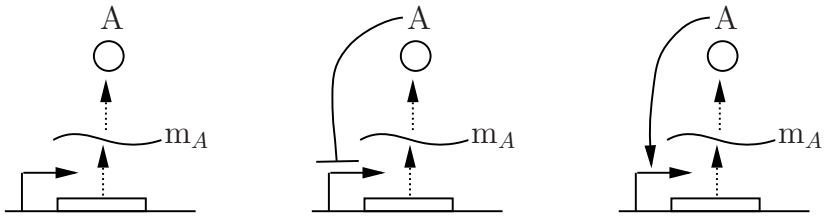
in which $[\text{Rep}]_{\text{tot}} = [\text{Rep}] + [\text{Rep:Ind}]$ is the total amount of repressor (bound and unbound to the inducer) and K_{d} is the dissociation constant of Ind binding to Rep. This expression of the repressor concentration needs to be substituted in the expression of the production rate $F([\text{Rep}])$.

Positive inducers can be handled similarly, except now we have to modify the binding reactions to only work in the presence of a regulatory protein bound to an inducer. For example, a positive inducer on an activator would have the modified reactions

$$\begin{aligned}
\text{Inducer binding:} &\quad \text{Act} + \text{Ind} \rightleftharpoons \text{Act:Ind} \\
\text{Activator binding:} &\quad \text{DNA}^{\text{p}} + \text{Act:Ind} \rightleftharpoons \text{DNA}^{\text{p}}\text{:Act:Ind} \\
\text{RNAP binding w/ activator:} &\quad \text{RNAP} + \text{DNA}^{\text{p}}\text{:Act:Ind} \rightleftharpoons \text{RNAP:DNA}^{\text{p}}\text{:Act:Ind} \\
\text{Isomerization:} &\quad \text{RNAP:DNA}^{\text{p}}\text{:Act:Ind} \longrightarrow \text{RNAP:DNA}^{\text{o}}\text{:Act:Ind} \\
\text{Start of transcription:} &\quad \text{RNAP:DNA}^{\text{o}}\text{:Act:Ind} \longrightarrow \text{RRNAP:DNA}^{\text{g,1}} + \\
&\qquad\qquad\qquad\qquad\qquad\qquad \text{DNA}^{\text{p}}\text{:Act:Ind}.
\end{aligned}$$

Hence, in the expression of the production rate $F([\text{Act}])$, we should substitute the concentration [Act:Ind] in place of [Act]. This concentration, in turn, is well approximated by its quasi-steady state value since binding reactions are much faster than isomerization and transcription.

Example 2.2 (Autoregulation of gene expression)**.** Consider the three circuits shown in Figure 2.18, representing a unregulated gene, a negatively autoregulated gene and a positively autoregulated gene. We want to model the dynamics of the protein A starting from zero initial conditions for the three different cases to understand how the three different circuit topologies affect dynamics.



(a) Unregulated (b) Negative autoregulation (c) Positive autoregulation

Figure 2.18: Autoregulation of gene expression. The three circuits control the expression of gene regulation using (a) unregulated, (b) negative autoregulation and (c) positive autoregulation.

The dynamics of the three circuits can be written in a common form,

$$\frac{dm_A}{dt} = F(A) - \delta m_A, \qquad \frac{dA}{dt} = \kappa m_A - \gamma A, \tag{2.19}$$

where $F(A)$ has the form

$$F_{\text{unreg}}(A) = \alpha_B, \qquad F_{\text{repress}}(A) = \frac{\alpha_B}{1+(A/K)^n} + \alpha_0, \qquad F_{\text{activate}}(A) = \frac{\alpha_A (A/K)^n}{1+(A/K)^n} + \alpha_B$$

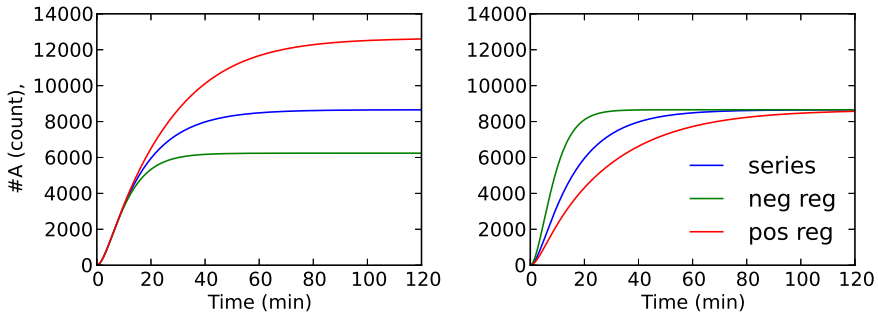
We choose the parameters to be

$$\begin{gathered} \alpha_A = 1/3, \qquad \alpha_B = 1/2, \qquad \alpha_0 = 5\times 10^{-4}, \\ \kappa = 20\log(2)/120, \qquad \delta = \log(2)/120, \qquad \gamma = \log(2)/600, \\ K = 10^4, \qquad n = 2, \end{gathered}$$

corresponding to biologically plausible values. Note that the parameters are chosen so that $f(0) \approx \alpha_B$ for each circuit.

Figure 2.19a shows the results of the simulation. We see that initial increase in protein concentration is identical for each circuit, consistent with our choice of Hill functions and parameters. As the expression level increases, the effects of positive and negative are seen, leading to different steady state expression levels. In particular, the negative feedback circuit reaches a lower steady state expression level while the positive feedback circuit settles to a higher value.

In some situations, it makes sense to ask whether different circuit topologies have different properties that might lead us to choose one over another. In the case where the circuit is going to be used as part of a more complex pathway, it may make the most sense to compare circuits that produce the same steady state concentration of the protein A. To do this, we must modify the parameters of the individual circuits, which can be done in a number of different ways: we can modify the promoter strengths, degradation rates, or other molecular mechanisms reflected in the parameters.



(a) (b)

Figure 2.19: Simulations for autoregulated gene expression. (a) Non-normalized expression levels. (b) Normalized expression.

The steady state expression level for the negative autoregulation case can be adjusted by using a stronger promoter (modeled by α_B) or ribosome binding site (modeled by κ). The equilibrium point for the negative autoregulation case is given by the solution of the equations

$$m_{A,e} = \frac{\alpha K^n}{\delta(K^n + A_e^n)}, \qquad A_e = \frac{\kappa}{\gamma} m_{A,e}.$$

These coupled equations can be solved for $m_{A,e}$ and A_e, but in this case we simply need to find values α'_B and κ' that give the same values as the unregulated case. For example, if we equate the mRNA levels of the unregulated system with that of the negatively autoregulated system, we have

$$\frac{\alpha_B}{\delta} = \frac{1}{\delta}\left(\frac{\alpha'_B K^n}{K^n + A_e^n} + \alpha_0\right) \quad \Longrightarrow \quad \alpha'_B = (\alpha_B - \alpha_0)\frac{K^n + A_e^n}{K^n}, \quad A_e = \frac{\alpha_B \kappa}{\delta\gamma},$$

where A_e is the desired equilibrium value (which we choose using the unregulated case as a guide).

A similar calculation can be done for the case of positive autoregulation, in this case decreasing the promoter parameters α_A and α_B so that the steady state values match. A simple way to do this is to leave α_A unchanged and decrease α_B to account for the positive feedback. Solving for α'_B to give the same mRNA levels as the unregulated case yields

$$\alpha'_B = \alpha_B - \alpha_A \frac{A_e^n}{K^n + A_e^n}.$$

Figure 2.19b shows simulations of the expression levels over time for the modified circuits. We see now that the expression levels all reach the same steady state value. The negative autoregulated circuit has the property that it reaches the steady

state more quickly, due to the increased rate of protein expression when A is small ($\alpha'_B > \alpha_B$). Conversely, the positive autoregulated circuit has a slower rate of expression than the constitutive case, since we have lowered the rate of protein expression when A is small. The initial higher and lower expression rates are compensated for via the autoregulation, resulting in the same expression level in steady state. ∇

We have described how the Hill function can model the regulation of a gene by a single transcription factor. However, genes can also be regulated by multiple transcription factors, some of which may be activators and some may be repressors. In this case, the promoter controlling the expression of the gene is called a combinatorial promoter. The mRNA production rate can thus take several forms depending on the roles (activators versus repressors) of the various transcription factors [3]. In general, the production rate resulting from a promoter that takes as input transcription factors $\mathrm{P_i}$ for $i \in \{1, ..., N\}$ will be denoted $F(P_1, ..., P_n)$.

Thus, the dynamics of a transcriptional module is often well captured by the ordinary differential equations

$$\frac{dm_{P_i}}{dt} = F(P_1, ..., P_n) - \delta_{P_i} m_{P_i}, \qquad \frac{dP_i}{dt} = \kappa_{P_i} m_{P_i} - \gamma_{P_i} P_i. \tag{2.20}$$

For a combinatorial promoter with two input proteins, an activator $\mathrm{P_a}$ and a repressor $\mathrm{P_r}$, in which the activator cannot bind if the repressor is bound to the promoter, the function $F(P_a, P_r)$ can be obtained by employing the competitive binding in the reduced order models of Section 2.1. In this case, assuming the activator has cooperativity n and the repressor has cooperativity m, we obtain the expression

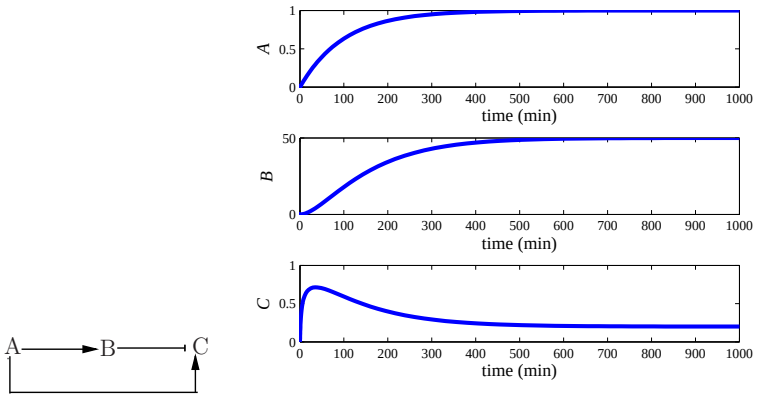
$$F(P_a, P_r) = \alpha \frac{(P_a/K_a)^n}{1 + (P_a/K_a)^n + (P_r/K_r)^m}. \tag{2.21}$$

Here, we have that $K_a = (K_{m,a} K_{\mathrm{d,a}})^{(1/n)}$, $K_r = (K_{m,r} K_{\mathrm{d,r}})^{(1/m)}$, in which $K_{\mathrm{d,a}}$ and $K_{\mathrm{d,r}}$ are the dissociation constants of the activator and repressor, respectively, from the DNA promoter site, while $K_{m,a}$ and $K_{m,r}$ are the dissociation constants for the cooperative binding reactions for the activator and repressor, respectively. In the case in which the activator is "leaky", that is, some transcription still occurs even when there is no activator, the above expression will be modified to

$$F(P_a, P_r) = \alpha \frac{(P_a/K_a)^n}{1 + (P_a/K_a)^n + (P_r/K_r)^m} + \alpha_0,$$

in which $\bar{\alpha}$ is the basal transcription rate when no activator is present. If such a basal rate can still be repressed by the repressor, the above expression modifies to the form

$$F(P_a, P_r) = \frac{\alpha (P_a/K_a)^n + \alpha_0}{1 + (P_a/K_a)^n + (P_r/K_r)^m}.$$



(a) (b)

Figure 2.20: The incoherent feed-forward loop (type I). (a) A schematic diagram of the circuit. (b) A simulation of a the model in equation (2.22) with $\alpha_A = 0.01$, $\gamma = 0.01$, $\alpha_B = 1$, $\alpha_C = 100$, $K_B = 0.001$, and $K_A = 1$.

Example 2.3 (Incoherent Feed-forward Loops)**.** Combinatorial promoters with two inputs are often used in systems where a logical "and" is required. As an example, we illustrate here an incoherent feed-forward loop (type I) [3]. Such a circuit is composed of three transcription factors A, B, and C, in which A directly activates C and B while B represses C. This is illustrated in Figure 2.20(a). This is different from a coherent feed-forward loop in which both A and B activate C. In the incoherent feed-forward loop, if we would like C to be high only when A is high and R is low ("and" gate), we can consider a combinatorial promoter in which the activator A and the repressor B competitively bind to the promoter of C. The resulting Hill function is given by the expression in (2.21). Depending on the values of the constants, the expression of C is low unless A is high and B is low. The resulting ODE model is given by the system

$$\begin{aligned} \frac{dA}{dt} &= \alpha_A - \gamma A \\ \frac{dB}{dt} &= \alpha_B \frac{A/K_A}{1+(A/K_A)} - \gamma B \\ \frac{dC}{dt} &= \alpha_C \frac{A/K_A}{1+(A/K_A)+(B/K_B)} - \gamma C \end{aligned} \tag{2.22}$$

in which we have assumed no cooperativity of binding for both the activator and the repressor. Upon a step of α_A, protein A binds to the promoter of C initiating transcription, so that protein C starts getting produced. At the same time, protein B is produced and accumulates until it reaches a large enough value to repress C. Hence, we can expect a pulse of C production for suitable parameter values. This is shown in Figure 2.20. In addition to being pulse generators, incoherent

feed-forward loops can accelerate the response time to step inputs, in which the response time is measured by the time the system takes to reach 90% of the steady state. ∇

Finally, a simple regulation mechanism is based on altering the half life of a protein. Specifically, the degradation rate of a protein is determined by the amounts of proteases present, which bind to recognition sites (degradation tags) and then degrade the protein. Degradation of a protein A by a protease P can then be modeled by the following two-step reaction

$$\mathrm{A} + \mathrm{P} \underset{d}{\overset{a}{\rightleftharpoons}} \mathrm{C} \xrightarrow{k} \mathrm{P},$$

in which C=AP is the complex of the protease Y bound to protein X. By the end of the reaction, protein X has been degraded to nothing, so that this reaction is often simplified to $\mathrm{X} \longrightarrow \emptyset$.

2.4 Post-Transcriptional Regulation

In addition to regulation of expression through modifications of the process of transcription, cells can also regulate the production and activity of proteins via a collection of other post-transcriptional modifications. These include methods of modulating the translation of proteins, as well as affecting the activity of a protein via changes in its conformation, as shown in Figure 2.9.

Allosteric modifications to proteins

In allosteric regulation, a regulatory molecule, called allosteric effector, binds to a site separate from the catalytic site (active site) of an enzyme. This binding causes a change in the three dimension conformation of the protein, turning off (or turning on) the catalytic site (Figure 2.21).

An allosteric effector can either be an activator or an inhibitor, just like inducers work for activation or inhibition of transcription factors. Inhibition can either be competitive or not competitive. In the case of competitive inhibition, the inhibitor competes with the substrate for binding the enzyme; that is, the substrate can bind to the enzyme only if the inhibitor is not bound. In the case of non-competitive inhibition, the substrate can be bound to the enzyme even if the latter is bound to the inhibitor. In this case, however, the product may not be able to form or may form at a lower rate, in which case, we have partial inhibition.

Activation can be absolute or not. Specifically, an activator is absolute when the enzyme can bind to the substrate only when bound to the activator. Otherwise, the activator is not absolute. In this section, we derive the expressions for the production rate of the active protein in an enzymatic reaction in the two most common

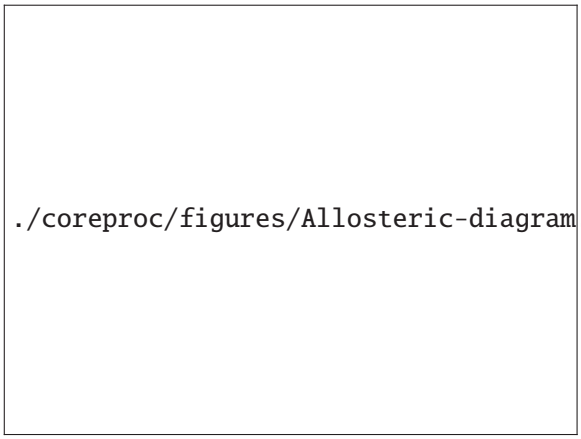


Figure 2.21: In allosteric regulation, a regulatory molecule binds to a site separate from the catalytic site (active site) of an enzyme. This binding causes a change in the three dimension conformation of the protein, turning off (or turning on) the catalytic site. Permission pending.

cases: when we have a (non-competitive) inhibitor I or an (absolute) activator A of the enzyme.

Allosteric inhibition

Consider the standard enzymatic reaction

$$\mathrm{E} + \mathrm{S} \underset{d}{\overset{a}{\rightleftharpoons}} \mathrm{ES} \xrightarrow{k} \mathrm{E} + \mathrm{P}$$

in which enzyme E binds to substrate S and transforms it into the product P. Let I be a (non-competitive) inhibitor of enzyme E so that when E is bound to I, the complex EI can still bind to substrate S, however, the complex EIS is non-productive, that is, it does not produce P. Then, we have the following additional reactions:

$$\mathrm{E} + \mathrm{I} \underset{k_-}{\overset{k_+}{\rightleftharpoons}} \mathrm{EI} \qquad \mathrm{ES} + \mathrm{I} \underset{k_-}{\overset{k_+}{\rightleftharpoons}} \mathrm{EIS} \qquad \mathrm{EI} + \mathrm{S} \underset{d}{\overset{a}{\rightleftharpoons}} \mathrm{EIS},$$

with the conservation laws (assuming S_{tot} is in much greater amounts than E_{tot})

$$E_{\mathrm{tot}} = E + [\mathrm{ES}] + [\mathrm{EI}] + [\mathrm{EIS}], \qquad S_{\mathrm{tot}} = S + P + [\mathrm{ES}] + [\mathrm{EIS}] \approx S + P.$$

The production rate of P is given by $dP/dt = k[\mathrm{ES}]$. Since binding reactions are very fast, we can assume all the complexes to be at the quasi-steady state. This gives

$$[\mathrm{EIS}] = \frac{a}{d}[\mathrm{EI}] \cdot S, \qquad [\mathrm{EI}] = \frac{k_+}{k_-} E \cdot I, \qquad [\mathrm{ES}] = \frac{1}{K_m} S \cdot E,$$

in which $K_m = (d+k)/a$ is the Michaelis-Menten constant. Using these expressions, the conservation law for the enzyme, and the fact that $a/d \approx 1/K_m$, we obtain

$$E = \frac{E_{\text{tot}}}{(I/K_{\text{d}}+1)(1+S/K_m)}, \text{ with } K_{\text{d}} = k_-/k_+,$$

so that

$$[\text{ES}] = \frac{S}{S+K_m}\frac{E_{\text{tot}}}{1+I/K_{\text{d}}}$$

and, as a consequence,

$$\frac{dP}{dt} = k_1 E_{\text{tot}}\left(\frac{1}{1+I/K_{\text{d}}}\right)\left(\frac{S}{S+K_m}\right).$$

Using the conservation law for S, this is also equivalent to

$$\frac{dP}{dt} = k_1 E_{\text{tot}}\left(\frac{1}{1+I/K_{\text{d}}}\right)\left(\frac{(S_{\text{tot}}-P)}{(S_{\text{tot}}-P)+K_m}\right).$$

In our earlier derivations of the Michaelis-Menten kinetics $V_{max} = k_1 E_{\text{tot}}$ was called the maximal speed of modification, which occurs when the enzyme is completely saturated by the substrate (Section 2.1). Hence, the effect of a non-competitive inhibitor is to decrease the maximal speed of modification by a factor $1/(1+I/K_{\text{d}})$.

Another type of inhibition occurs when the inhibitor is competitive, that is, when I is bound to E, the complex EI cannot bind to protein S. Since E can either bind to I or S (not both), I competes against S for binding to E. See Exercise 2.10.

Allosteric activation

In this case, the enzyme E can transform S to its active form only when it is bound to A. Also, we assume that E cannot bind S unless E is bound to A (from here, the name absolute activator). The reactions are therefore modified to be

$$\text{E} + \text{A} \underset{k_-}{\overset{k_+}{\rightleftharpoons}} \text{EA}$$

and

$$\text{EA} + \text{S} \underset{d}{\overset{a}{\rightleftharpoons}} \text{EAS} \overset{k}{\rightarrow} \text{P} + \text{EA},$$

with conservation laws

$$E_{\text{tot}} = E + [\text{EA}] + [\text{EAS}], \quad S_{\text{tot}} \approx S + P.$$

The production rate of P is given by $dP/dt = k[\text{EAS}]$. Assuming as above that the complexes are at the quasi-steady state, we have that

$$[\text{EA}] = \frac{E \cdot A}{K_{\text{d}}}, \qquad [\text{EAS}] = \frac{S \cdot [\text{EA}]}{K_m},$$

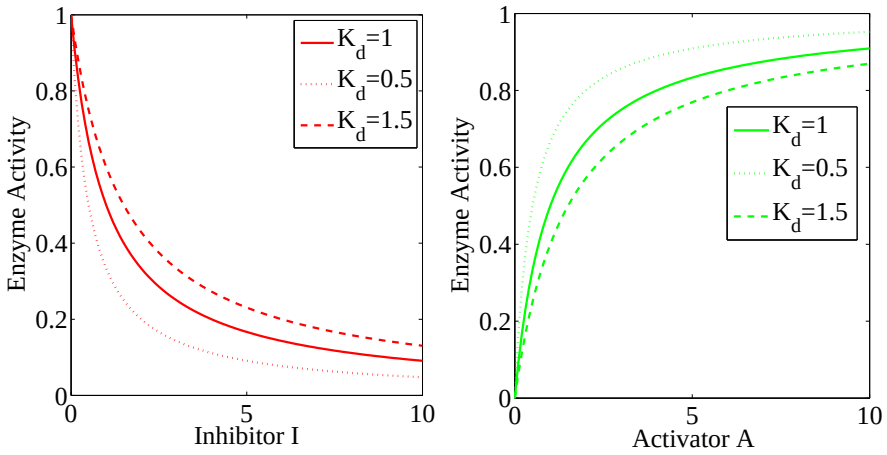


Figure 2.22: Enzyme activity in the presence of allosteric effectors (activators or inhibitors). The red plots show the enzyme activity in the presence of an inhibitor as a function of the inhibitor concentration. The green plots show the enzyme activity in the presence of an activator as a function of the activator concentration. The different plots show the effect of the dissociation constant.

which, using the conservation law for E, leads to

$$E = \frac{E_{\text{tot}}}{(1+S/K_m)(1+A/K_{\text{d}})} \quad \text{and} \quad [\text{EAS}] = \left(\frac{A}{A+K_{\text{d}}}\right)\left(\frac{S}{S+K_m}\right)E_{\text{tot}}.$$

Hence, we have that

$$\frac{dP}{dt} = kE_{\text{tot}}\left(\frac{A}{A+K_{\text{d}}}\right)\left(\frac{S}{S+K_m}\right).$$

Using the conservation law for S, this is also equivalent to

$$\frac{dP}{dt} = kE_{\text{tot}}\left(\frac{A}{A+K_{\text{d}}}\right)\left(\frac{(S_{\text{tot}}-P)}{(S_{\text{tot}}-P)+K_m}\right).$$

The effect of an absolute activator is to modulate the maximal speed of modification by a factor $A/(A+K_{\text{d}})$.

Figure 2.22 shows the behavior of the enzyme activity as a function of the allosteric effector. As the dissociation constant decreases, that is, the affinity of the effector increases, a very small amount of effector will cause the enzyme activity to be completely “on” in the case of the activator and completely “off” in the case of the inhibitor.

Another type of activation occurs when the activator is not absolute, that is, when E can bind to S directly, but cannot activate S unless the complex ES first binds A (see Exercise 2.11).

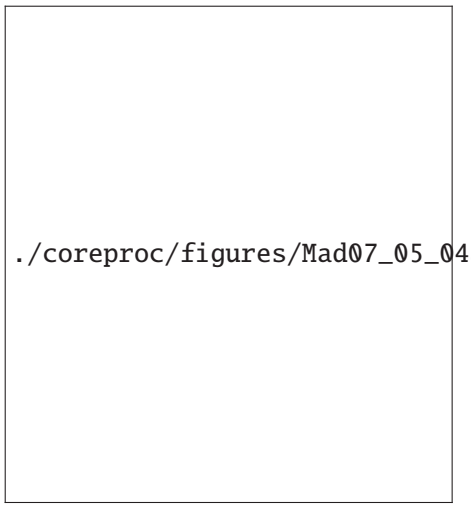


Figure 2.23: Phosphorylation of a protein via a kinase. Reproduced from Madhani [58]; permission pending.

Covalent modifications to proteins

In addition to regulation that controls transcription of DNA into mRNA, a variety of mechanisms are available for controlling expression after mRNA is produced. These include control of splicing and transport from the nucleus (in eukaryotes), the use of various secondary structure patterns in mRNA that can interfere with ribosomal binding or cleave the mRNA into multiple pieces, and targeted degradation of mRNA. Once the polypeptide chain is formed, additional mechanisms are available that regulate the folding of the protein as well as its shape and activity level. We briefly describe some of the major mechanisms here.

Review

Material to be written: sRNA, riboswitches.

One of the most common types of post-transcriptional regulation is through the *phosphorylation* of proteins. Phosphorylation is an enzymatic process in which a phosphate group is added to a protein and the resulting conformation of the protein changes, usually from an inactive configuration to an active one. The enzyme that adds the phosphate group is called a *kinase* (or sometimes a *phosphotransferase*) and it operates by transferring a phosphate group from a bound ATP molecule to the protein, leaving behind ADP and the phosphorylated protein. *Dephosphorylation* is a complementary enzymatic process that can remove a phosphate group from a protein. The enzyme that performs dephosphorylation is called a *phosphatase*. Figure 2.23 shows the process of phosphorylation in more detail.

Phosphorylation is often used as a regulatory mechanism, with the phosphorylated version of the protein being the active conformation. Since phosphorylation and dephosphorylation can occur much more quickly than protein production and

degradation, it is used in biological circuits in which a rapid response is required. One common pattern is that a signaling protein will bind to a ligand and the resulting allosteric change allows the signaling protein to serve as a kinase. The newly active kinase then phosphorylates a second protein, which modulates other functions in the cell. Phosphorylation cascades can also be used to amplify the effect of the original signal; we will describe this in more detail in Section 2.5.

Kinases in cells are usually very specific to a given protein, allowing detailed signaling networks to be constructed. Phosphatases, on the other hand, are much less specific, and a given phosphatase species may desphosphorylate many different types of proteins. The combined action of kinases and phosphatases is important in signaling since the only way to deactivate a phosphorylated protein is by removing the phosphate group. Thus phosphatases are constantly "turning off" proteins, and the protein is activated only when sufficient kinase activity is present.

Phosphorylation of a protein occurs by the addition of a charged phosphate (PO_4) group to the serine (Ser), threonine (Thr) or tyrosine (Tyr) amino acids. Similar covalent modifications can occur by the attachment of other chemical groups to select amino acids. *Methylation* occurs when a methyl group (CH_3) is added to lysine (Lys) and is used for modulation of receptor activity and in modifying histones that are used in chromatin structures. *Acetylation* occurs when an acetyl group ($COCH_3$) is added to lysine and is also used to modify histones. *Ubiquitination* refers to the addition of a small protein, ubiquitin, to lysine; the addition of a polyubiquitin chain to a protein targets it for degradation.

Covalent modification is a post-translational protein modification that affects the activity of the protein. It plays an important role both in the control of metabolism and in signal transduction. Here, we focus on *reversible* cycles of modification, in which a protein is interconverted between two forms that differ in activity either because of effects on the kinetics relative to substrates or for altered sensitivity to effectors.

At a high level, a covalent modification cycle involves a target protein X, an enzyme Z for modifying it, and a second enzyme Y for reversing the modification (see Figure 2.24). We call X^* the activated protein. There are often allosteric effectors or further covalent modification systems that regulate the activity of the modifying enzymes, but we do not consider this added level of complexity here. There are several types of covalent modification, depending on the type of activation of the protein. *Phosphorylation* is a covalent modification that takes place mainly in eukaryotes and involves activation of the inactive protein X by addition of a phosphate group, PO_4. In this case, the enzyme Z is called a *kinase* while the enzyme Y is called *phosphatase*. Another type of covalent modification, which is very common in both procaryotes and eukaryotes, is *methylation*. Here, the inactive protein is activated by the addition of a methyl group, CH_3.

The reactions describing this system are given by the following two enzymatic

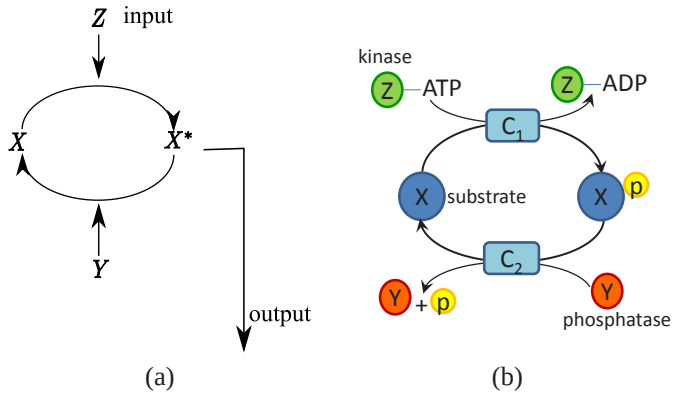


Figure 2.24: (Left) General diagram representing a covalent modification cycle. (Right) Detailed view of a phoshorylation cycle including ATP, ADP, and the exchange og the phosphate group "p".

reactions, also called a two step reaction model,

$$\mathrm{Z} + \mathrm{X} \underset{d_1}{\overset{a_1}{\rightleftharpoons}} \mathrm{C}_1 \xrightarrow{k_1} \mathrm{X}^* + \mathrm{Z}, \qquad \mathrm{Y} + \mathrm{X}^* \underset{d_2}{\overset{a_2}{\rightleftharpoons}} \mathrm{C}_2 \xrightarrow{k_2} \mathrm{X} + \mathrm{Y},$$

in which we have let C_1=ZX be the kinase/protein complex and C_2=X^*Y be the active protein/phosphatase complex. The corresponding ODE model is given by

$$\begin{aligned}
\frac{dZ}{dt} &= -a_1 Z \cdot X + (k_1 + d_1) C_1, & \frac{dX^*}{dt} &= k_1 C_1 - a_2 Y \cdot X^* + d_2 C_2, \\
\frac{dX}{dt} &= -a_1 Z \cdot X + d_1 C_1 + k_2 C_2, & \frac{dC_2}{dt} &= a_2 Y \cdot X^* - (d_2 + k_2) C_2, \\
\frac{dC_1}{dt} &= a_1 Z \cdot X - (d_1 + k_1) C_1, & \frac{dY}{dt} &= -a_2 Y \cdot X^* + (d_2 + k_2) C_2.
\end{aligned}$$

Furthermore, we have that the total amounts of enzymes Z and Y are conserved. Denote the total concentrations of Z and Y by Z_{tot}, Y_{tot}, respectively. Then, we have also the conservation laws $Z + C_1 = Z_{\mathrm{tot}}$ and $Y + C_2 = Y_{\mathrm{tot}}$. We can thus reduce the above system of ODE to the following one, in which we have substituted $Z = Z_{\mathrm{tot}} - C_1$ and $Y = Y_{\mathrm{tot}} - C_2$:

$$\begin{aligned}
\frac{dC_1}{dt} &= a_1 (Z_{\mathrm{tot}} - C_1) \cdot X - (d_1 + k_1) C_1, \\
\frac{dX^*}{dt} &= k_1 C_1 - a_2 (Y_{\mathrm{tot}} - C_2) \cdot X^* + d_2 C_2, \\
\frac{dC_2}{dt} &= a_2 (Y_{\mathrm{tot}} - C_2) \cdot X^* - (d_2 + k_2) C_2.
\end{aligned}$$

As for the case of the enzymatic reaction, this system cannot be analytically integrated. To simplify it, we can perform a similar approximation as done for the

enzymatic reaction. In particular, the complexes C_1 and C_2 are often assumed to reach their steady state values very quickly because $a_1, d_1, a_2, d_2 \gg k_1, k_2$. Therefore, we can approximate the above system by substituting for C_1 and C_2 their steady state values, given by the solutions to

$$a_1(Z_{\text{tot}} - C_1) \cdot X - (d_1 + k_1)C_1 = 0$$

and

$$a_2(Y_{\text{tot}} - C_2) \cdot X^* - (d_2 + k_2)C_2 = 0.$$

By solving these equations, we obtain that

$$C_2 = \frac{Y_{\text{tot}} X^*}{X^* + K_{m,2}}, \quad \text{with} \quad K_{m,2} = \frac{d_2 + k_2}{a_2}$$

and

$$C_1 = \frac{Z_{\text{tot}} X}{X + K_{m,1}}, \quad \text{with} \quad K_{m,1} = \frac{d_1 + k_1}{a_1}.$$

As a consequence, the ODE model of the phosphorylation system can be well approximated by

$$\frac{dX^*}{dt} = k_1 \frac{Z_{\text{tot}} X}{X + K_{m,1}} - a_2 \frac{Y_{\text{tot}} K_{m,2}}{X^* + K_{m,2}} \cdot X^* + d_2 \frac{Y_{\text{tot}} X^*}{X^* + K_{m,2}},$$

which, considering that $a_2 K_{m,2} - d_2 = k_2$, leads finally to

$$\frac{dX^*}{dt} = k_1 \frac{Z_{\text{tot}} X}{X + K_{m,1}} - k_2 \frac{Y_{\text{tot}} X^*}{X^* + K_{m,2}}. \tag{2.23}$$

We will come back to the modeling of this system after we have introduced singular perturbation theory, through which we will be able to perform a formal analysis and mathematically characterize the assumptions needed for approximating the original system by the first order ODE model (2.23). In the model of equation (2.23), we have that $X = X_{\text{tot}} - X^* - C_1 - C_2$ by the conservation laws. A standard assumption is that the amounts of enzymes are small compared to the amount of substrate, so that $X \approx X_{\text{tot}} - X^*$ [33].

Ultrasensitivity

One relevant aspect of the response of the covalent modification cycle to its input is the sensitivity of the steady state characteristic curve. Specifically, what parameters affect the shape of the steady state response is a crucial question. To determine the steady state characteristics, which shows how the steady state of X^* changes when the input stimulus Z_{tot} is changed, we set $dX^*/dt = 0$ in equation (2.23). Using the

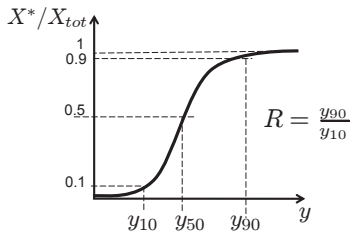


Figure 2.25: Steady state characteristic curve showing the relevance of the response coefficient for ultrasensitivity. As $R \to 1$, the points y_{10} and y_{90} tend to each other.

approximation $X \approx X_{\text{tot}} - X^*$, denoting $V_1 := k_1 Z_{\text{tot}}$, $V_2 := k_2 Y_{\text{tot}}$, $\bar{K}_1 := K_{m,1}/X_{\text{tot}}$, and $\bar{K}_2 := K_{m,2}/X_{\text{tot}}$, we obtain

$$y := \frac{V_1}{V_2} = \frac{X^*/X_{\text{tot}}\left(\bar{K}_1 + (1 - X^*/X_{\text{tot}})\right)}{(\bar{K}_2 + X^*/X_{\text{tot}})(1 - X^*/X_{\text{tot}})}. \tag{2.24}$$

We are interested in the shape of the steady state curve of X^* as function of y. This shape is usually characterized by two key parameters: the response coefficient, denoted R, and the point of half maximal induction, denoted y_{50}. Let y_α denote the value of y corresponding to having X^* equal α% of the maximum value of X^* obtained for $y = \infty$, which is equal to X_{tot}. Then, the response coefficient is defined as

$$R := \frac{y_{90}}{y_{10}},$$

and measures how switch-like the response is (Figure 2.25). When $R \to 1$ the response becomes switch-like. In the case in which the steady state characteristic is a Hill function, we have that $X^* = (y/K)^n/(1 + (y/K)^n)$, so that $y_\alpha = (\alpha/(100 - \alpha))^{(1/n)}$ and as a consequence

$$R = (81)^{(1/n)}, \text{ or equivalently } n = \frac{\log(81)}{\log(R)}.$$

Hence, when $n = 1$, that is, the characteristic is of the Michaelis-Menten type, we have that $R = 81$, while when n increases, R decreases. Usually, when $n > 1$ the response is referred to as *ultrasensitive*. The formula $n = \log(81)/\log(R)$ is often employed to estimate the *apparent Hill coefficient* of a dose response curve (the input/output steady state characteristic curve obtained from experimental data) since R can be calculated for any response curve directly from the data points.

In the case of the current system, from equation (2.24), we have that

$$y_{90} = \frac{(\bar{K}_1 + 0.1)\ 0.9}{(\bar{K}_2 + 0.9)\ 0.1} \quad \text{and} \quad y_{10} = \frac{(\bar{K}_1 + 0.9)\ 0.1}{(\bar{K}_2 + 0.1)\ 0.9},$$

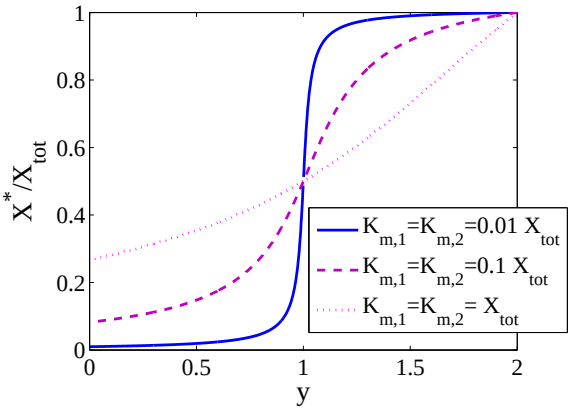


Figure 2.26: Steady state characteristics of a covalent modification cycle as a function of the Michaelis-Menten constants $K_{m,1}$ and $K_{m,2}$.

so that

$$R = 81\frac{(\bar{K_1}+0.1)(\bar{K_2}+0.1)}{(\bar{K_2}+0.9)(\bar{K_1}+0.9)}. \tag{2.25}$$

As a consequence, when $\bar{K_1}, \bar{K_2} \gg 1$, we have that $R \to 81$, which gives a Michaelis-Menten type of response. If instead $\bar{K_1}, \bar{K_2} \ll 0.1$, we have that $R \to 1$, which corresponds to a theoretic Hill coefficient $n \gg 1$, that is, a switch-like response (Figure 2.26). In particular, if we have, for example, $\bar{K_1} = \bar{K_2} = 10^{-2}$, we obtain an apparent Hill coefficient grater than 13. This type of ultrasensitivity is usually referred to as *zero-order ultrasensitivity*. The reason of this name is due to the fact that when $K_{m,1}$ is much smaller than the amount of protein substrate X, we have that $Z_{tot}X/(K_{m,1}+X) \approx Z_{tot}$. Hence, the forward modification rate is "zero order" in the substrate concentration (no free enzyme is left, all is bound to the substrate).

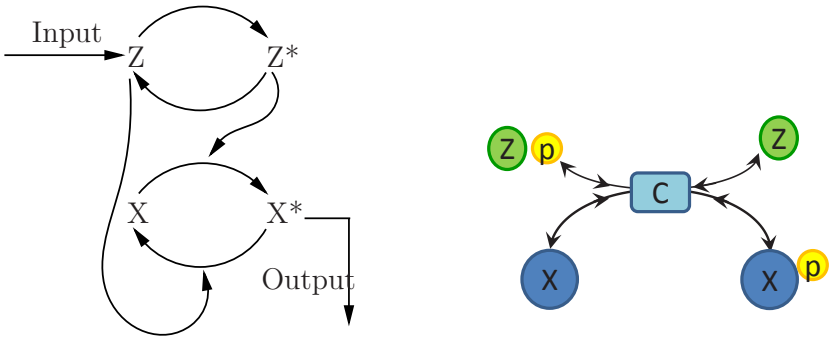
One can study the behavior also of the point of half maximal induction

$$y_{50} = \frac{\bar{K_1}+0.5}{\bar{K_2}+0.5},$$

to find that as $\bar{K_2}$ increases, it decreases and that as $\bar{K_1}$ increases, it increases.

Phosphotransfer systems

Phosphotransfer systems are also a common motif in cellular signal transduction. These structures are composed of proteins that can phosphorylate each other. In contrast to kinase-mediated phosphorylation, where the phosphate donor is usually ATP, in phosphotransfer the phosphate group comes from the donor protein itself (Figure 2.27). Each protein carrying a phosphate group can donate it to the next protein in the system through a reversible reaction. In this section, we describe a module extracted from the phosphotransferase system [87].



(a) (b)

Figure 2.27: (a) Diagram of a phosphotransfer system. (b) Proteins X and Z are transferring the phosphate group p to each other.

Let X be a transcription factor in its inactive form and let X^* be the same transcription factor once it has been activated by the addition of a phosphate group. Let Z^* be a phosphate donor, that is, a protein that can transfer its phosphate group to the acceptor X. The standard phosphotransfer reactions [78] can be modeled according to the two-step reaction model

$$Z^* + X \underset{k_2}{\overset{k_1}{\rightleftharpoons}} C_1 \underset{k_4}{\overset{k_3}{\rightleftharpoons}} X^* + Z,$$

in which C_1 is the complex of Z bound to X bound to the phosphate group. Additionally, protein Z can be phosphorylated and protein X^* dephosphorylated by other phosphotransfer interactions. These reactions are modeled as one step reactions depending only on the concentrations of Z and X^*, that is,

$$Z \xrightarrow{\pi_1} Z^*, \qquad X^* \xrightarrow{\pi_2} X.$$

Protein X is assumed to be conserved in the system, that is, $X_{\text{tot}} = X + C_1 + X^*$. We assume that protein Z is produced with time-varying production rate $k(t)$ and decays with rate γ. The ODE model corresponding to this system is thus given by the equations

$$\begin{aligned}
\frac{dZ}{dt} &= k(t) - \gamma Z + k_3 C_1 - k_4 X^* Z - \pi_1 Z \\
\frac{dC_1}{dt} &= k_1 X_{\text{tot}} \left(1 - \frac{X^*}{X_{\text{tot}}} - \frac{C_1}{X_{\text{tot}}}\right) Z^* - k_3 C_1 - k_2 C_1 + k_4 X^* Z \\
\frac{dZ^*}{dt} &= \pi_1 Z + k_2 C_1 - k_1 X_{\text{tot}} \left(1 - \frac{X^*}{X_{\text{tot}}} - \frac{C_1}{X_{\text{tot}}}\right) Z^* \\
\frac{dX^*}{dt} &= k_3 C_1 - k_4 X^* Z - \pi_2 X^*.
\end{aligned} \tag{2.26}$$

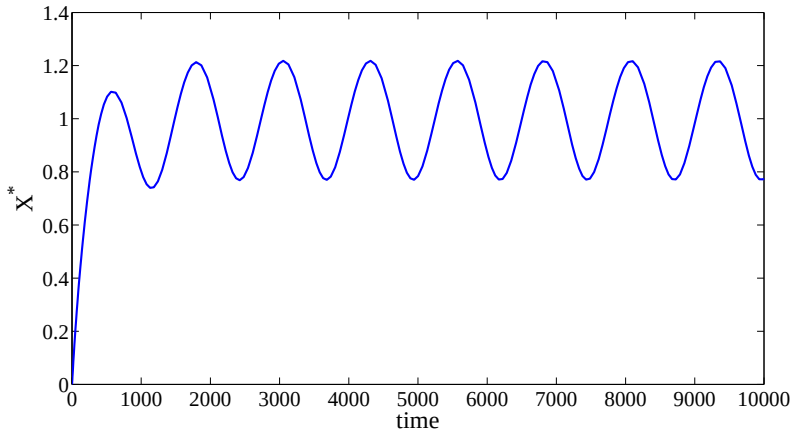


Figure 2.28: Output response of the phosphotransfer system with a step signal $k(t) = 1 + 0.5\sin(\omega t)$. The parameters are given by $\gamma = 0.01$, $X_{\text{tot}} = 5000$, $k_1 = k_2 = k_3 = k_4 = \pi_1 = \pi_2 = 0.01$.

Sample simulation results when the input is a time-varying (periodic) stimulus are shown in Figure 2.28. The output X^* well "tracks" the input stimulus by virtue of the fast phosphotransfer reactions.

This model will be considered again in Chapter 7 when the phosphotransfer system is proposed as a possible realization of an insulation device to buffer systems from retroactivity effects.

2.5 Cellular Subsystems

In the previous section we have studied how to model a variety of core processes that occure in cells. In this section we consider a few common "subsystems" in which these processes are combined for specific purposes.

Intercellular signaling: MAPK cascades

The Mitogen Activated Protein Kinase (MAPK) cascade is a recurrent structural motif in several signal transduction pathways (Figure 2.29). The cascade consists of a MAPK kinase kinase (MAPKKK), denoted X_0, a MAPK kinase (MAPKK), denoted X_1, and a MAPK, denoted X_2. MAPKKKs activate MAPKKs by phosphorylation at two conserved sites and MAPKKs activate MAPKs by also phosphorylation at conserved sites. The cascade relays signals from the plasma membrane to targets in the cytoplasm and nucleus. It has been extensively studied and modeled. Here, we provide two different models. First, we build a modular model by viewing the system as the composition of single phosphorylation cycle modules

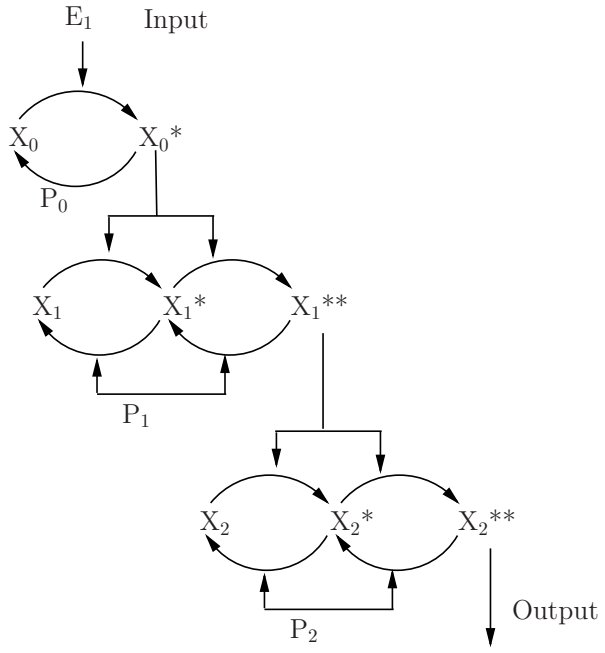


Figure 2.29: Schematic representing the MAPK cascade. It has three levels: the first one has a single phosphorylation, while the second and the third ones have a double phosphorylation.

(whose ODE model was derived earlier) and double phosphorylation cycle modules, whose ODE model we derive here. Then, we provide the full list of reactions describing the cascade and construct a mechanistic ODE model from scratch. We will then highlight the difference between the two derived models.

Double phosphorylation model. Consider the double phosphorylation motif in Figure 2.30. The reactions describing the system are given by

$$\mathrm{E}_1 + \mathrm{X} \underset{d_1}{\overset{a_1}{\rightleftharpoons}} \mathrm{C}_1 \xrightarrow{k_1} \mathrm{X}^* + \mathrm{E}_1, \qquad \mathrm{E}_2 + \mathrm{X}^* \underset{d_2}{\overset{a_2}{\rightleftharpoons}} \mathrm{C}_2 \xrightarrow{k_2} \mathrm{X} + \mathrm{E}_2,$$
$$\mathrm{X}^* + \mathrm{E}_1 \underset{d_1^*}{\overset{a_1^*}{\rightleftharpoons}} \mathrm{C}_3 \xrightarrow{k_1^*} \mathrm{X}^{**} + \mathrm{E}_1, \qquad \mathrm{E}_2 + \mathrm{X}^{**} \underset{d_2^*}{\overset{a_2^*}{\rightleftharpoons}} \mathrm{C}_4 \xrightarrow{k_2^*} \mathrm{X}^* + \mathrm{E}_2,$$

in which C_1 is the complex of E_1 with X, C_2 is the complex of E_2 with X^*, C_3 is the complex of E_1 with X^*, and C_4 is the complex of E_2 with X^{**}. The conservation laws are given by

$$E_1 + C_1 + C_3 = E_{1,\mathrm{tot}}, \qquad E_2 + C_2 + C_4 = E_{2,\mathrm{tot}},$$
$$X_{\mathrm{tot}} = X + X^* + X^{**} + C_1 + C_2 + C_3 + C_4 \approx X + X^* + X^{**},$$

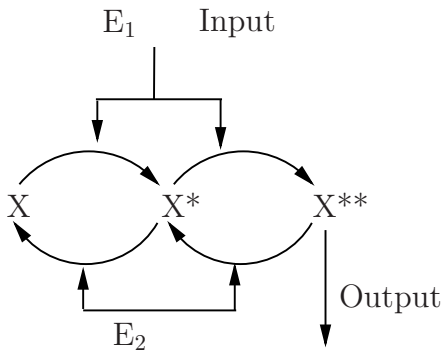


Figure 2.30: Schematic representing a double phosphorylation cycle. E_1 is the input and X^{**} is the output.

in which we have assumed the the total amounts of enzymes are small compared to the total amount of substrate as we have explained earlier. As performed earlier, we assume that the complexes are at the quasi-steady state since binding reactions are very fast compared to the catalytic rates. This gives the Michaelis-Menten form for the amount of formed complexes:

$$C_1 = E_{1,\text{tot}} \frac{K_1^* X}{K_1^* X + K_1 X^* + K_1 K_1^*}, \qquad C_3 = E_{1,\text{tot}} \frac{K_1 X^*}{K_1^* X + K_1 X^* + K_1 K_1^*},$$
$$C_2 = E_{2,\text{tot}} \frac{K_2^* X^*}{K_2^* X^* + K_2 X^{**} + K_2 K_2^*}, \qquad C_4 = E_{2,\text{tot}} \frac{K_2 X^{**}}{K_2^* X^* + K_2 X^{**} + K_2 K_2^*},$$

in which $K_i = (d_i + k_i)/a_i$ and $K_i^* = (d_i^* + k_i^*)/a_i^*$ are the Michaelis-Menten constants for the enzymatic reactions. Since the complexes are at the quasi-steady state, it follows that

$$\frac{d}{dt} X^* = k_1 C_1 - k_2 C_2 - k_1^* C_3 + k_2^* C_4,$$
$$\frac{d}{dt} X^{**} = k_1^* C_3 - k_2^* C_4,$$

from which, substituting the expressions of the complexes, we obtain that

$$\frac{d}{dt} X^* = E_{1,\text{tot}} \frac{k_1 X K_1^* - k_1^* X^* K_1}{K_1^* X + K_1 X^* + K_1^* K_1} + E_{2,\text{tot}} \frac{k_2^* X^{**} K_2 - k_2 X^* K_2^*}{K_2^* X^* + K_2 X^{**} + K_2 K_2^*}$$
$$\frac{d}{dt} X^{**} = k_1^* E_{1,\text{tot}} \frac{K_1 X^*}{K_1^* X + K_1 X^* + K_1 K_1^*} - k_2^* E_{2,\text{tot}} \frac{K_2 X^{**}}{K_2^* X^* + K_2 X^{**} + K_2 K_2^*},$$

in which $X = X_{\text{tot}} - X^* - X^{**}$.

Modular model of MAPK cascades

In this section, to simplify notation, we denote "MAPK" by X_2. In a modular composition framework, the output of one stage becomes an input to the next stage downstream of it. Hence, X_0^* becomes the input enzyme that activates the phosphorylation of X_1, and X_1^{**} becomes the input enzyme that activates the phosphorylation of X_2. Let $(a_{1,i}, d_{1,i}, k_{1,i})$ and $(a_{2,i}, d_{2,i}, k_{2,i})$ be the association, dissociation, and catalytic rates for the forward and backward enzymatic reactions, respectively, for the first cycle at stage $i \in \{0,1,2\}$. Similarly, let $(a^*_{1,i}, d^*_{1,i}, k^*_{1,i})$ and $(a^*_{2,i}, d^*_{2,i}, k^*_{2,i})$ be the association, dissociation, and catalytic rates for the forward and backward enzymatic reactions, respectively, for the second cycle at stage $i \in \{1,2\}$. Also, denote by $K_{1,i}$ and $K_{2,i}$ for $i \in \{0,1,2\}$ the Michaelis-Menten constants of the forward and backward enzymatic reactions, respectively, of the first cycle at stage i. Similarly, denote $K^*_{1,i}$ and $K^*_{2,i}$ for $i \in \{1,2\}$ be the Michaelis-Menten constants of the forward and backward enzymatic reactions, respectively, of the second cycle at stage i. Let $P_{1,\text{tot}}$ and $P_{2,\text{tot}}$ be the total amounts of the X_1 and X_2 phosphatases, respectively. Then, the modular ODE model of the MAPK cascade is given by

$$
\begin{aligned}
\frac{d}{dt}X_0^* &= k_{1,0}E_{1,\text{tot}}\frac{X_0}{X_0+K_{1,0}} - k_{2,0}P_{0,\text{tot}}\frac{X_0^*}{X_0^*+K_{2,0}}\\
\frac{d}{dt}X_1^* &= X_0^*\frac{k_{1,1}\ X_0\ K^*_{1,1} - k^*_{1,1}\ X_1^*\ K_{1,1}}{K^*_{1,1}\ X_1+K_{1,1}\ X_1^*+K_{1,1}K^*_{1,1}} + P_{1,\text{tot}}\frac{k^*_{2,1}\ K_{2,1}\ X_1^{**} - k_{2,1}\ X_1^*\ K^*_{2,1}}{K^*_{2,1}\ X_1^*+K_{2,1}\ X_1^{**}+K_{2,1}K^*_{2,1}}\\
\frac{d}{dt}X_1^{**} &= k^*_{1,1}\ X_0^*\frac{X_1^*\ K_{1,1}}{K^*_{1,1}\ X_1+K_{1,1}\ X_1^*+K_{1,1}K^*_{1,1}} - k^*_{2,1}\ P_{1,\text{tot}}\frac{X_1^{**}\ K_{2,1}}{K^*_{2,1}\ X_1^*+K_{2,1}\ X_1^{**}+K_{2,1}K^*_{2,1}}\\
\frac{d}{dt}X_2^* &= X_1^{**}\frac{k_{1,2}X_2\ K^*_{1,2} - k^*_{1,2}\ X_2^*\ K_{1,2}}{K^*_{1,2}\ X_2+K_{1,2}\ X_2^*+K^*_{1,2}\ K_{1,2}} + P_{2,\text{tot}}\frac{k^*_{2,2}\ K_{2,2}\ X_2^{**} - k_{2,2}\ X_2^*\ K^*_{2,2}}{K^*_{2,2}\ X_2^*+K_{2,2}\ X_2^{**}+K_{2,2}\ K^*_{2,2}}\\
\frac{d}{dt}X_2^{**} &= k^*_{1,2}\ X_1^{**}\frac{X_2^*\ K_{1,2}}{K^*_{1,2}\ X_2+K_{1,2}\ X_2^*+K^*_{1,2}\ K_{1,2}} - k^*_{2,2}\ P_{2,\text{tot}}\frac{X_2^{**}\ K_{2,2}}{K^*_{2,2}\ X_2^*+K_{2,2}\ X_2^{**}+K_{2,2}\ K^*_{2,2}}
\end{aligned}
\tag{2.27}
$$

in which, letting $X_{0,\text{tot}}, X_{1,tot}$ and $X_{2,\text{tot}}$ represent the total amounts of each stage protein, we have $X_0 = X_{0,\text{tot}} - X_0^*$, $X_1 = X_{1,\text{tot}} - X_1^* - X_1^{**}$ and $X_2 = X_{2,\text{tot}} - X_2^* - X_2^{**}$.

Mechanistic model of the MAPK cascade

We now give the entire set of reactions for the MAPK cascade of Figure 2.29 as they are found in standard references (Huang-Ferrell model [41]):

$$\mathrm{E}_1 + \mathrm{X}_0 \underset{d_{1,0}}{\overset{a_{1,0}}{\rightleftharpoons}} \mathrm{C}_1 \xrightarrow{k_{1,0}} \mathrm{X}_0^* + \mathrm{E}_1 \qquad \mathrm{P}_0 + \mathrm{X}_0^* \underset{d_{2,0}}{\overset{a_{2,0}}{\rightleftharpoons}} \mathrm{C}_2 \xrightarrow{k_{2,0}} \mathrm{X}_0 + \mathrm{P}_0$$

$$\mathrm{X}_0^* + \mathrm{X}_1 \underset{d_{1,1}}{\overset{a_{1,1}}{\rightleftharpoons}} \mathrm{C}_3 \xrightarrow{k_{1,1}} \mathrm{X}_1^* + \mathrm{X}_0^* \qquad \mathrm{X}_1^* + \mathrm{P}_1 \underset{d_{2,1}}{\overset{a_{2,1}}{\rightleftharpoons}} \mathrm{C}_4 \xrightarrow{k_{2,1}} \mathrm{X}_1 + \mathrm{P}_1$$

$$\mathrm{X}_0^* + \mathrm{X}_1^* \underset{d_{1,1}^*}{\overset{a_{1,1}^*}{\rightleftharpoons}} \mathrm{C}_5 \xrightarrow{k_{1,1}^*} \mathrm{X}_1^{**} + \mathrm{X}_0^* \qquad \mathrm{X}_1^{**} + \mathrm{P}_1 \underset{d_{2,1}^*}{\overset{a_{2,1}^*}{\rightleftharpoons}} \mathrm{C}_6 \xrightarrow{k_{2,1}^*} \mathrm{X}_1^* + \mathrm{P}_1$$

$$\mathrm{X}_1^{**} + \mathrm{X}_2 \underset{d_{1,2}}{\overset{a_{1,2}}{\rightleftharpoons}} \mathrm{C}_7 \xrightarrow{k_{1,2}} \mathrm{X}_2^* + \mathrm{X}_1^{**} \qquad \mathrm{X}_2^* + \mathrm{P}_2 \underset{d_{2,2}}{\overset{a_{2,2}}{\rightleftharpoons}} \mathrm{C}_8 \xrightarrow{k_{2,2}} \mathrm{X}_2 + \mathrm{P}_2$$

$$\mathrm{X}_1^{**} + \mathrm{X}_2^* \underset{d_{1,2}^*}{\overset{a_{1,2}^*}{\rightleftharpoons}} \mathrm{C}_9 \xrightarrow{k_{1,2}^*} \mathrm{X}_2^{**} + \mathrm{X}_1^{**} \qquad \mathrm{X}_2^{**} + \mathrm{P}_2 \underset{d_{2,2}^*}{\overset{a_{2,2}^*}{\rightleftharpoons}} \mathrm{C}_{10} \xrightarrow{k_{2,2}^*} \mathrm{X}_2^* + \mathrm{P}_2,$$

with conservation laws

$$\begin{aligned}
X_{0,\mathrm{tot}} &= X_0 + X_0^* + C_1 + C_2 + C_3 + C_5 \\
X_{1,\mathrm{tot}} &= X_1 + X_1^* + C_3 + X_1^{**} + C_4 + C_5 + C_6 + C_7 + C_9 \\
X_{2,\mathrm{tot}} &= X_2 + X_2^* + X_2^{**} + C_7 + C_8 + C_9 + C_{10} \\
E_{1,\mathrm{tot}} &= E_1 + C_1, \quad P_{0,\mathrm{tot}} = P_0 + C_2 \\
P_{1,\mathrm{tot}} &= P_1 + C_4 + C_6 \\
P_{2,\mathrm{tot}} &= P_2 + C_8 + C_{10}.
\end{aligned}$$

The corresponding ODE model is given by

$$
\begin{aligned}
\frac{d}{dt}C_1 &= a_{1,0}E_1\ X_0 - (d_{1,0}+k_{1,0})\ C_1\\
\frac{d}{dt}X_0^* &= k_{1,0}\ C_1 + d_{2,0}\ C_2 - a_{2,0}\ P_0\ X_0^* + (d_{1,1}+k_{1,1})\ C_3 - a_{1,1}\ X_1\ X_0^*\\
&\quad + (d_{1,1}^*+k_{1,1}^*)\ C_5 - a_{1,1}^*\ X_0^*\ X_1^*\\
\frac{d}{dt}C_2 &= a_{2,0}\ P_0\ X_0^* - (d_{2,0}+k_{2,0})\ C_2\\
\frac{d}{dt}C_3 &= a_{1,1}\ X_1\ X_0^* - (d_{1,1}+k_{1,1})\ C_3\\
\frac{d}{dt}X_1^* &= k_{1,1}\ C_3 + d_{2,1}\ C_4 - a_{2,1}\ X_1^*\ P_1 + d_{1,1}^* C_5 - a_{1,1}^*\ X_1^*\ X_0^* + k_{2,1}^*\ C_6\\
\frac{d}{dt}C_4 &= a_{2,1}\ X_1^*\ P_1 - (d_{2,1}+k_{2,1})\ C_4\\
\frac{d}{dt}C_5 &= a_{1,1}^*\ X_0^*\ X_1^* - (d_{1,1}^*+k_{1,1}^*)\ C_5\\
\frac{d}{dt}X_1^{**} &= k_{1,1}^*\ C_5 - a_{2,1}^*\ X_1^*\ P_1 + d_{2,1}^*\ C_6 - a_{1,2}\ X_1^{**}\ X_2\\
&\quad + (d_{1,2}+k_{1,2})\ C_7 - a_{1,2}^*\ X_1^{**}\ X_2^* + (d_{1,2}^*+k_{1,2}^*)\ C_9\\
\frac{d}{dt}C_6 &= a_{2,1}^*\ X_1^{**}\ P_1 - (d_{2,1}^*+k_{2,1}^*)\ C_6\\
\frac{d}{dt}C_7 &= a_{1,2}^*\ X_1^*\ X_2 - (d_{1,2}^*+k_{1,2}^*)\ C_7\\
\frac{d}{dt}X_2^* &= -a_{2,2}\ X_2^*\ P_2 + d_{2,2}\ C_8 - a_{1,2}^*\ X_2^*\ X_2^{**} + d_{1,2}^*\ C_9 + C_{10}\ K_{10}\\
\frac{d}{dt}C_8 &= a_{2,2}^*\ X_2^*\ P_2 - (d_{2,2}+k_{2,2})\ C_8\\
\frac{d}{dt}X_2^{**} &= k_{1,2}^*\ C_9 - a_{2,2}^*\ X_2^{**}\ P_2 + d_{2,2}^*\ C_{10}\\
\frac{d}{dt}C_9 &= a_{1,2}^*\ X_1^{**}\ X_2^* - (d_{1,2}^*+k_{1,2}^*)\ C_9\\
\frac{d}{dt}C_{10} &= a_{2,2}^*\ X_2^{**}\ P_2 - (d_{2,2}^*+k_{2,2}^*)\ C_{10}.
\end{aligned}
$$

Assuming as before that the total amounts of enzymes are much smaller than the total amounts of substrates ($E_{1,\mathrm{tot}}, P_{0,\mathrm{tot}}, P_{1,\mathrm{tot}}, P_{2,\mathrm{tot}} \ll X_{0,\mathrm{tot}}, X_{1,\mathrm{tot}}, X_{2,\mathrm{tot}}$), we can approximate the conservation laws as

$$
\begin{aligned}
X_{0,\mathrm{tot}} &\approx X_0 + X_0^* + C_3 + C_5,\\
X_{1,\mathrm{tot}} &\approx X_1 + X_1^* + C_3 + X_1^{**} + C_5 + C_7 + C_9,\\
X_{2,\mathrm{tot}} &\approx X_2 + X_2^* + X_2^{**} + C_7 + C_9.
\end{aligned}
$$

Using these and assuming that the complexes are at the quasi-steady state, we obtain the following functional dependencies:

$$C_1 = f_1(X_0^*, X_1^*, X_1^{**}, X_2^*, X_2^{**}), \qquad C_2 = f_2(X_0^*),$$
$$C_3 = f_3(X_0^*, X_1^*, X_1^{**}, X_2^*, X_2^{**}), \qquad C_5 = f_5(X_0^*, X_1^*),$$
$$C_7 = f_7(X_1^*, X_1^{**}, X_2^*, X_2^{**}), \qquad C_9 = f_9(X_1^{**}, X_2^*).$$

The fact that C_7 depends on X_2^* and X_2^{**} illustrates that the dynamics of the second stage are influenced by those of the third stage. Similarly, the fact that C_3 depends on $X_1^*, X_1^{**}, X_2^*, X_2^{**}$ indicates that the dynamics of the first stage are influenced by those of the second stage and by that of the third stage. The phenomenon by which the behavior of a "module" is influenced by that of its downstream clients is called *retroactivity*, which is a phenomenon similar to impedance in electrical systems and to back-effect in mechanical systems. It will be studied at length in Chapter 7.

This fact is in clear contrast with the ODE model obtained by modular composition, in which each stage dynamics depended upon the variables of the upstream stages and not upon those of the downstream stages. That is, from equations (2.27), it is apparent that the dynamics of X_0^* (first stage) do not depend on the variables of the second stage (X_1, X_1^*, X_1^{**}). In turn, the dynamics of X_1^* and X_1^{**} (second stage) do not depend on the variables of the third stage (X_2^* and X_2^{**}). Indeed modular composition does not consider the fact that the proteins of each stage are "used-up" in the process of transmitting information to the downstream stages. This backward effect has been theoretically shown to lead to sustained oscillations in the MAPK cascade [76]. By contrast, the modular ODE model of MAPK cascades does not give rise to sustained oscillations.

Properties of the MAPK Cascade

The stimulus-response curve obtained with the mechanistic model predicts that the response of the MAPKKK to the stimulus $E_{1,\text{tot}}$ is of the Michaelis-Menten type. By contrast, the stimulus-response curve obtained for the MAPKK and MAPK are sigmoidal and show high Hill coefficients, which increases from the MAPKK response to the MAPK response. That is, an increase ultrasensitivity is observed moving down in the cascade (Figure 2.31). These model observations persist when key parameters, such as the Michaelis-Menten constants are changed [41]. Furthermore, zero-order ultrasensitivity effects can be observed. Specifically, if the amounts of MAPKK were increased, one would observe a higher apparent Hill coefficient for the response of MAPK. Similarly, if the values of the K_m for the reactions in which the MAPKK takes place were decreased, one would also observe a higher apparent Hill coefficient for the response of MAPK. Double phosphorylation is also key to obtain a high apparent Hill coefficient. In fact, a cascade in which the double phosphorylation was assumed to occur through a one-step model

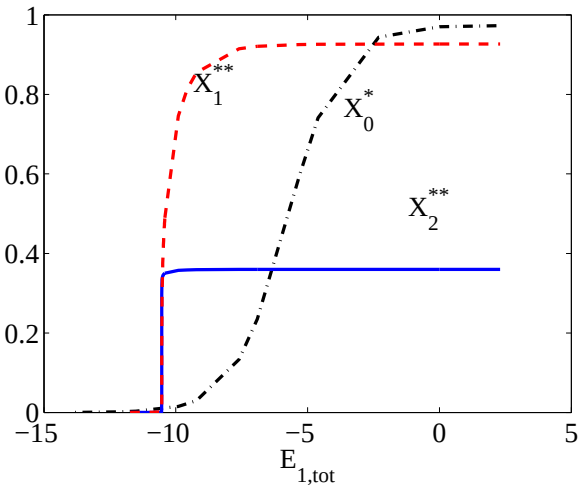


Figure 2.31: Dose response of the MAPK cascade for every stage. Simulations from the model of [76].

(similar to single phosphorylation) predicted substantially lower apparent Hill coefficients.

Review

Additional topics to be added later:

1. Transport across the membrane
2. Membrane receptors, ligand binding, G-proteins

Exercises

2.1 (BE 150, Winter 2011) Consider a cascade of three activators X→Y→ Z. Protein X is initially present in the cell in its inactive form. The input signal of X, S_x, appears at time t=0. As a result, X rapidly becomes active and binds the promoter of gene Y, so that protein Y starts to be produced at rate β. When Y levels exceed a threshold K, gene Z begins to be transcribed and translated at rate β. All proteins have the same degradation/dilution rate γ.

(a) What are the concentrations of proteins Y and Z as a function of time?

(b) What is the minimum duration of the pulse S_x such that Z will be produced?

(c) What is response time of protein Z with respect to the time of addition of S_x?

2.2 (Hill function for a cooperative repressor) Consider a repressor that binds to an operator site as a dimer:

$$\begin{aligned}
&\text{R1:} \quad \mathrm{R} + \mathrm{R} \rightleftharpoons \mathrm{R}_2 \\
&\text{R2:} \quad \mathrm{R}_2 + \mathrm{DNA}^{\mathrm{p}} \rightleftharpoons \mathrm{R}_2{:}\mathrm{DNA} \\
&\text{R3:} \quad \mathrm{RNAP} + \mathrm{DNA}^{\mathrm{p}} \rightleftharpoons \mathrm{RNAP{:}DNA}^{\mathrm{p}}
\end{aligned}$$

Assume that the reactions are at equilibrium and that the RNA polymerase concentration is large (so that [RNAP] is roughly constant). Show that the ratio of the concentration of $\mathrm{RNA{:}DNA}^{\mathrm{p}}$ to the total amount of DNA, D_{tot}, can be written as a Hill function

$$f(R) = \frac{[\mathrm{RNAP{:}DNA}]}{D_{\mathrm{tot}}} = \frac{\alpha}{K + R^2}$$

and give expressions for α and K.

2.3 (Switch-like behavior in cooperative binding) Derive the expressions of C and A at the steady state when you have the cooperative binding reactions

$$\mathrm{B} + \mathrm{B} + \ldots + \mathrm{B} \underset{k_2}{\overset{k_1}{\rightleftharpoons}} \mathrm{B_n}, \quad \mathrm{B_n} + \mathrm{A} \underset{d}{\overset{a}{\rightleftharpoons}} \mathrm{C}, \quad \text{and} \quad A + C = A_{\mathrm{tot}}.$$

Make MATLAB plots of the expressions that you obtain and verify that as n increases the functions become more switch-like.

2.4 Consider the following modification of the competitive binding reactions:

$$\mathrm{B_a} + \mathrm{A} \underset{d}{\overset{a}{\rightleftharpoons}} \mathrm{C}, \; \mathrm{B_r} + \mathrm{A} \underset{\bar{d}}{\overset{\bar{a}}{\rightleftharpoons}} \bar{C},$$

and

$$\mathrm{C} + \mathrm{B_r} \underset{d'}{\overset{a'}{\rightleftharpoons}} \mathrm{C}', \text{ and } \bar{C} + \mathrm{B_a} \underset{\bar{d}'}{\overset{\bar{a}'}{\rightleftharpoons}} \mathrm{C}'$$

with $A_{\mathrm{tot}} = A + C + \bar{C} + C'$. What are the steady state expressions for A and C? What information do you deduce from these expressions if A is a promoter, Ba is an activator protein, and C is the activator/DNA complex that makes the gene transcriptionally active?

2.5 Consider the case of a competitive binding of an activator A and a repressor R with D and assume that before they can bind D they have to cooperatively bind according to the following reactions:

$$\mathrm{A} + \mathrm{A} + \ldots + \mathrm{A} \underset{k_2}{\overset{k_1}{\rightleftharpoons}} \mathrm{A_n}, \qquad \mathrm{R} + \mathrm{R} + \ldots + \mathrm{R} \underset{\bar{k}_2}{\overset{\bar{k}_1}{\rightleftharpoons}} \mathrm{R_m},$$

in which the complex A_n contains n molecules of A and the complex R_m contains m molecules of R. The competitive binding reactions with A are given by

$$A_n + D \underset{d}{\overset{a}{\rightleftharpoons}} C,\ R_m + D \underset{d'}{\overset{a'}{\rightleftharpoons}} C',$$

and $D_{tot} = D + C + C'$. What are the steady state expressions for C and D?

2.6 Assume that we have an activator B_a and a repressor protein B_r. We want to obtain an input function such that when a lot of B_a is present, the gene is transcriptionally active only if there is no B_r, when low amounts of B_a are present, the gene is transcriptionally inactive (with or without B_r). Write down the reactions among B_a, B_r, and complexes with the DNA (A) that lead to such an input function. Demonstrate that indeed the set of reactions you picked leads to the desired input function.

2.7 Consider the phosphorylation reactions described in Section 2.4, but suppose that the kinase concentration Z is not constant, but is produced and decays according to the reaction $Z \underset{k(t)}{\overset{\gamma}{\rightleftharpoons}} \emptyset$. How should the system in equation (2.23) be modified? Use a MATLAB simulation to apply a periodic input stimulus $k(t)$ using parameter values: $k_1 = k_2 = 1$, $a_1 = a_2 = d_1 = d_2 = 10$, $\gamma = 0.01$. Is the cycle capable of "tracking" the input stimulus? If yes, to what extent? What are the tracking properties depending on?

2.8 Another model for the phosphorylation reactions, referred to as one step reaction model, is given by $Z + X \longrightarrow X^* + Z$ and $Y + X^* \longrightarrow X + Y$, in which the complex formations are neglected. Write down the ODE model and comparing the differential equation of X^* to that of equation (2.23), list the assumptions under which the one step reaction model is a good approximation of the two step reaction model.

2.9 (Transcriptional regulation with delay) Consider a repressor or activator B^* modeled by a Hill function $F_(B)$. Show that in the presence of transcriptional delay τ^m, the dynamics of the active mRNA can be written as

$$\frac{dm^*(t)}{dt} = e^{-\tau^m} F(B(t-\tau^m)) - \bar{\delta} m^*.$$

2.10 (Competitive Inhibition) Derive the expression of the production rate of X^* in the phosphorylation cycle in the presence of a competitive inhibitor I.

2.11 (Non-absolute activator) Derive the expression of the production rate of X^* in the phosphorylation cycle in the presence of a non-absolute activator A.

Chapter 3
Analysis of Dynamic Behavior

In this chapter, we describe some of the tools from dynamical systems and feedback control theory that will be used in the rest of the text to analyze and design biological circuits, building on tools already described in AM08. We focus here on deterministic models and the associated analyses; stochastic methods are given in Chapter 4.

Prerequisites. Readers should have a understanding of the tools for analyzing stability of solutions to ordinary differential equations, at the level of Chapter 4 of AM08. We will also make use of linearized input/output models in state space, based on the techniques described in Chapter 5 of AM08 and the frequency domain techniques described in Chapters 8–10.

3.1 Analysis Near Equilibria

As in the case of many other classes of dynamical systems, a great deal of insight into the behavior of a biological system can be obtained by analyzing the dynamics of the system subject to small perturbations around a known solution. We begin by considering the dynamics of the system near an equilibrium point, which is one of the simplest cases and provides a rich set of methods and tools.

In this section we will model the dynamics of our system using the input/output modeling formalism described in Chapter 1:

$$\dot{x} = f(x, \theta, u), \qquad y = h(x, \theta), \tag{3.1}$$

where $x \in \mathbb{R}^n$ is the system state, $\theta \in \mathbb{R}^p$ are the system parameters and $u \in \mathbb{R}^q$ is a set of external inputs (including disturbances and noise). The system state x is a vector whose components will represent concentration of species, such as proteins, kinases, DNA promoter sites, inducers, allosteric effectors, etc. The system parameters θ is also a vector, whose components will represent biochemical parameters such as association and dissociation rates, production rates, decay rates, dissociation constants, etc. The input u is a vector whose components will represent a number of possible physical entities, including the concentration of transcription factors, DNA concentration, kinases concentration, etc. The output $y \in \mathbb{R}^m$ of the system represents quantities that can be measured or that are used to interconnect subsystem models to form larger models.

Example 3.1 (Transcriptional component)**.** Consider a promoter controlling a gene g that can be regulated by a transcription factor Z. Let m_G and G represent the mRNA and protein expressed by gene g. This system can be viewed as a system, in which $u = Z$ is the concentration of transcription factor regulating the promoter, the state $x = (x_1, x_2)$ is such that $x_1 = m_G$ is the concentration of mRNA and $x_2 = G$ is the concentration of protein, and $y = G = x_2$ is the concentration of protein G. Assuming that the transcription factor regulating the promoter is a repressor, the system dynamics can be described by the following system

$$\frac{dx_1}{dt} = \frac{\alpha}{1+(u/K)^n} - \delta x_1, \qquad \frac{dx_2}{dt} = \kappa x_1 - \gamma x_2, \qquad y = x_2 \tag{3.2}$$

in which $\theta = (\alpha, K, \delta, \kappa, \gamma, n)$ is the vector of system parameters. In this case, we have that

$$f(x,\theta,u) = \begin{pmatrix} \dfrac{\alpha}{1+(u/K)^n} - \delta x_1 \\ \kappa x_1 - \gamma x_2 \end{pmatrix}, \qquad h(x,\theta) = x_2.$$

∇

Note that we have chosen to explicitly model the system parameters θ, which can be thought of as an additional set of (mainly constant) inputs to the system.

Equilibrium points and stability [AM08]

We begin by considering the case where the input u and parameters θ in equation (3.1) are fixed and hence we can write the dynamics of the system as

$$\frac{dx}{dt} = f(x). \tag{3.3}$$

An *equilibrium point* of a dynamical system represents a stationary condition for the dynamics. We say that a state x_e is an equilibrium point for a dynamical system if $f(x_e) = 0$. If a dynamical system has an initial condition $x(0) = x_e$, then it will stay at the equilibrium point: $x(t) = x_e$ for all $t \geq 0$.

Equilibrium points are one of the most important features of a dynamical system since they define the states corresponding to constant operating conditions. A dynamical system can have zero, one or more equilibrium points.

The *stability* of an equilibrium point determines whether or not solutions nearby the equilibrium point remain close, get closer or move further away. An equilibrium point x_e is *stable* if solutions that start near x_e stay close to x_e. Formally, we say that the equilibrium point x_e is stable if for all $\epsilon > 0$, there exists a $\delta > 0$ such that

$$\|x(0) - x_e\| < \delta \quad \Longrightarrow \quad \|x(t) - x_e\| < \epsilon \quad \text{for all } t > 0,$$

where $x(t)$ represents the solution the the differential equation (3.3) with initial condition $x(0)$. Note that this definition does not imply that $x(t)$ approaches x_e as

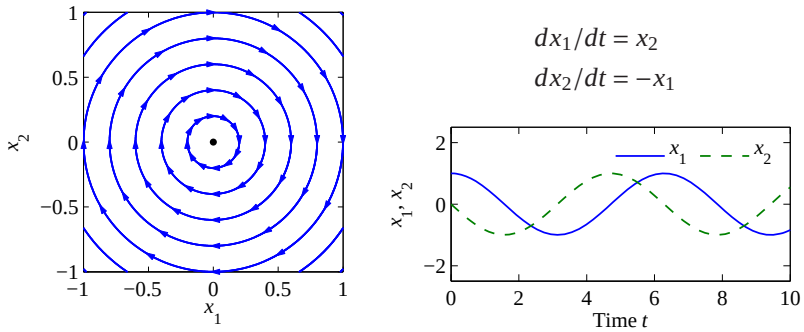


Figure 3.1: Phase portrait (trajectories in the state space) on the left and time domain simulation on the right for a system with a single stable equilibrium point. The equilibrium point x_e at the origin is stable since all trajectories that start near x_e stay near x_e.

time increases but just that it stays nearby. Furthermore, the value of δ may depend on ϵ, so that if we wish to stay very close to the solution, we may have to start very, very close ($\delta \ll \epsilon$). This type of stability, which is illustrated in Figure 3.1, is also called *stability in the sense of Lyapunov*. If an equilibrium point is stable in this sense and the trajectories do not converge, we say that the equilibrium point is *neutrally stable*.

An example of a neutrally stable equilibrium point is shown in Figure 3.1. From the phase portrait, we see that if we start near the equilibrium point, then we stay near the equilibrium point. Indeed, for this example, given any ϵ that defines the range of possible initial conditions, we can simply choose $\delta = \epsilon$ to satisfy the definition of stability since the trajectories are perfect circles.

An equilibrium point x_e is *asymptotically stable* if it is stable in the sense of Lyapunov and also $x(t) \to x_e$ as $t \to \infty$ for $x(0)$ sufficiently close to x_e. This corresponds to the case where all nearby trajectories converge to the stable solution for large time. Figure 3.2 shows an example of an asymptotically stable equilibrium point.

Note from the phase portraits that not only do all trajectories stay near the equilibrium point at the origin, but that they also all approach the origin as t gets large (the directions of the arrows on the phase portrait show the direction in which the trajectories move).

An equilibrium point x_e is *unstable* if it is not stable. More specifically, we say that an equilibrium point x_e is unstable if given some $\epsilon > 0$, there does *not* exist a $\delta > 0$ such that if $\|x(0) - x_e\| < \delta$, then $\|x(t) - x_e\| < \epsilon$ for all t. An example of an unstable equilibrium point is shown in Figure 3.3.

The definitions above are given without careful description of their domain of applicability. More formally, we define an equilibrium point to be *locally stable* (or *locally asymptotically stable*) if it is stable for all initial conditions $x \in B_r(a)$,

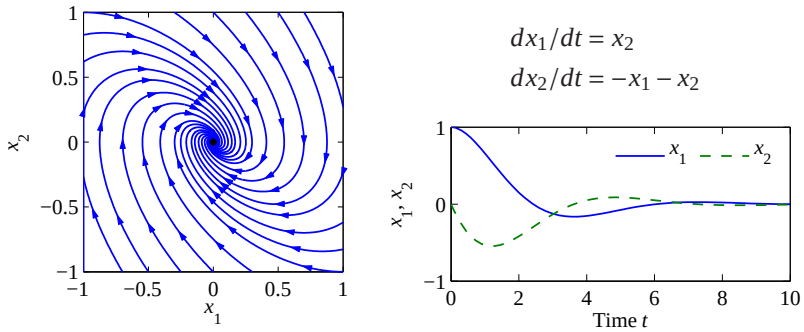


Figure 3.2: Phase portrait and time domain simulation for a system with a single asymptotically stable equilibrium point. The equilibrium point x_e at the origin is asymptotically stable since the trajectories converge to this point as $t \to \infty$.

where

$$B_r(a) = \{x : \|x - a\| < r\}$$

is a ball of radius r around a and $r > 0$. A system is *globally stable* if it is stable for all $r > 0$. Systems whose equilibrium points are only locally stable can have interesting behavior away from equilibrium points, as we explore in the next section.

To better understand the dynamics of the system, we can examine the set of all initial conditions that converge to a given asymptotically stable equilibrium point. This set is called the *region of attraction* for the equilibrium point. In general, computing regions of attraction is difficult. However, even if we cannot determine the region of attraction, we can often obtain patches around the stable equilibria that are attracting. This gives partial information about the behavior of the system.

For planar dynamical systems, equilibrium points have been assigned names based on their stability type. An asymptotically stable equilibrium point is called a *sink* or sometimes an *attractor*. An unstable equilibrium point can be either a *source*, if all trajectories lead away from the equilibrium point, or a *saddle*, if some trajectories lead to the equilibrium point and others move away (this is the situation pictured in Figure 3.3). Finally, an equilibrium point that is stable but not asymptotically stable (i.e., neutrally stable, such as the one in Figure 3.1) is called a *center*.

Example 3.2 (Bistable gene circuit)**.** Consider a system composed of two genes that express transcription factors that repress each other as shown in Figure 3.4. Denoting the concentration of protein A by x_1 and that of protein B by x_2 and neglecting the mRNA dynamics, the system can be modeled by the following differential equations:

$$\frac{dx_1}{dt} = \frac{\alpha_1}{1 + (x_2/K_2)^n} - \gamma x_1, \qquad \frac{dx_2}{dt} = \frac{\alpha_2}{1 + (x_1/K_1)^n} - \gamma x_2.$$

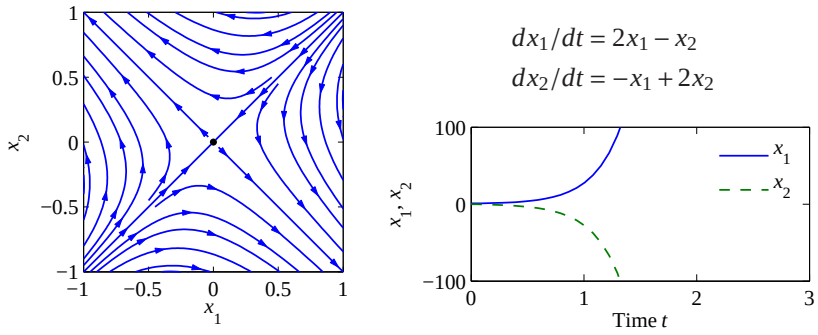


Figure 3.3: Phase portrait and time domain simulation for a system with a single unstable equilibrium point. The equilibrium point x_e at the origin is unstable since not all trajectories that start near x_e stay near x_e. The sample trajectory on the right shows that the trajectories very quickly depart from zero.

Figure 3.4(b) shows the phase portrait of the system. This system is bi-stable because there are two (asymptotically) stable equilibria. Specifically, the trajectories converge to either of two possible equilibria: one where x_1 is high and x_2 is low and the other where x_1 is low and x_2 is high. A trajectory will approach the first one if the initial condition is below the dashed line, called the separatrix, while it will approach the second one if the initial condition is above the separatrix. Hence, the region of attraction of the first equilibrium is the region of the plane below the separatrix and the region of attraction of the second one is the portion of the plane above the separatrix. ∇

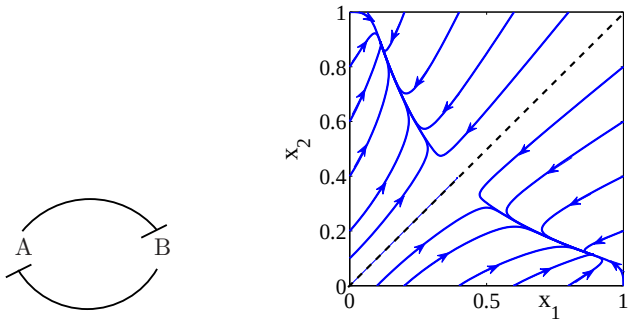
Nullcline Analysis

Nullcline analysis is a simple and intuitive way to determine the stability of an equilibrium point for systems in $\mathbb{R}^2$. Consider the system with $x = (x_1, x_2) \in \mathbb{R}^2$ described by the differential equations

$$\frac{dx_1}{dt} = f_1(x_1, x_2), \qquad \frac{dx_2}{dt} = f_2(x_1, x_2).$$

The nullclines of this system are given by the two curves in the x_1, x_2 plane in which $f_1(x_1, x_2) = 0$ and $f_2(x_1, x_2) = 0$. The nullclines intersect at the equilibria of the system x_e. Figure 3.5 shows an example in which there is a unique equilibrium.

The stability of the equilibrium is deduced by inspecting the direction of the trajectory of the system starting at initial conditions x close to the equilibrium x_e. The direction of the trajectory can be obtained by determining the signs of f_1 and f_2 in each of the regions in which the nullclines partition the plane around the equilibrium x_e. If $f_1 < 0$ ($f_1 > 0$), we have that x_1 is going to decrease (increase) and similarly if $f_2 < 0$ ($f_2 > 0$), we have that x_2 is going to decrease (increase). In



(a) (b)

Figure 3.4: (a) Diagram of a bistable gene circuit composed of two genes. (b) Phase plot showing the trajectories converging to either one of the two possible stable equilibria depending on the initial condition. The parameters are $\alpha_1 = \alpha_2 = 1$, $K_1 = K_2 = 0.1$, and $\gamma = 1$.

Figure 3.5, we show a case in which $f_1 < 0$ on the right-hand side of the nullcline $f_1 = 0$ and $f_1 > 0$ on the left-hand side of the same nullcline. Similarly, we have chosen a case in which $f_2 < 0$ above the nullcline $f_2 = 0$ and $f_2 > 0$ below the same nullcline. Given these signs, it is clear (see the figure) that starting from any point x close to x_e the vector field will always point toward the equilibrium x_e and hence the trajectory will tend toward such equilibrium. In this case, it then follows that the equilibrium x_e is asymptotically stable.

Example 3.3 (Negative autoregulation)**.** As an example, consider expression of a gene with negative feedback. Let x_1 represent the mRNA concentration and x_2 represent the protein concentration. Then, a simple model (in which for simplicity we have assumed all parameters to be 1) is given by

$$\frac{dx_1}{dt} = \frac{1}{1+x_2} - x_1, \qquad \frac{dx_2}{dt} = x_1 - x_2,$$

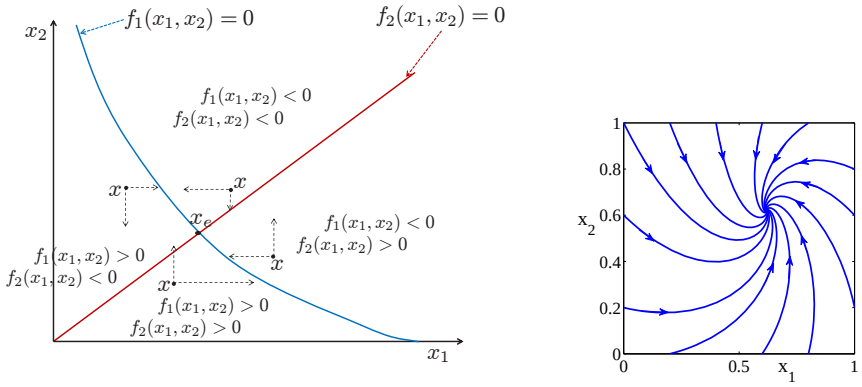
so that $f_1(x_1, x_2) = 1/(1+x_2) - x_1$ and $f_2(x_1, x_2) = x_1 - x_2$. Figure 3.5(a) exactly represents the situation for this example. In fact, we have that

$$f_1(x_1, x_2) < 0 \quad \Longleftrightarrow \quad x_1 > \frac{1}{1+x_2}, \qquad f_2(x_1, x_2) < 0 \quad \Longleftrightarrow \quad x_2 > x_1,$$

which provides the direction of the vector field as shown in Figure 3.5. As a consequence, the equilibrium point is stable. The phase plot of Figure 3.5(b) confirms this fact since the trajectories all converge to the unique equilibrium point. ∇

Stability analysis via linearization

For systems with more than two states, the graphical technique of nullcline analysis cannot be used. Hence, we must resort to other techniques to determine stability.



(a) (b)

Figure 3.5: (a) Example of nullclines for a system with a single equilibrium point x_e. To understand the stability of the equilibrium point x_e, one traces the direction of the vector field (f_1, f_2) in each of the four regions in which the nullcline partition the plane. If in each region the vector field points toward the equilibrium point, then such a point is asymptotically stable. (b) Phase plot diagram for the negative autoregulation example.

Consider a linear dynamical system of the form

$$\frac{dx}{dt} = Ax, \quad x(0) = x_0, \tag{3.4}$$

where $A \in \mathbb{R}^{n \times n}$. For a linear system, the stability of the equilibrium at the origin can be determined from the eigenvalues of the matrix A:

$$\lambda(A) = \{s \in \mathbb{C} : \det(sI - A) = 0\}.$$

The polynomial $\det(sI - A)$ is the *characteristic polynomial* and the eigenvalues are its roots. We use the notation λ_j for the jth eigenvalue of A and $\lambda(A)$ for the set of all eigenvalues of A, so that $\lambda_j \in \lambda(A)$. For each eigenvalue λ_j there is a corresponding eigenvector $v_j \in \mathbb{R}^n$, which satisfies the equation $Av_j = \lambda_j v_j$.

In general λ can be complex-valued, although if A is real-valued, then for any eigenvalue λ, its complex conjugate λ^* will also be an eigenvalue. The origin is always an equilibrium point for a linear system. Since the stability of a linear system depends only on the matrix A, we find that stability is a property of the system. For a linear system we can therefore talk about the stability of the system rather than the stability of a particular solution or equilibrium point.

The easiest class of linear systems to analyze are those whose system matrices are in diagonal form. In this case, the dynamics have the form

$$\frac{dx}{dt} = \begin{pmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{pmatrix} x. \tag{3.5}$$

It is easy to see that the state trajectories for this system are independent of each other, so that we can write the solution in terms of n individual systems $\dot{x}_j = \lambda_j x_j$. Each of these scalar solutions is of the form

$$x_j(t) = e^{\lambda_j t} x_j(0).$$

We see that the equilibrium point $x_e = 0$ is stable if $\lambda_j \leq 0$ and asymptotically stable if $\lambda_j < 0$.

Another simple case is when the dynamics are in the block diagonal form

$$\frac{dx}{dt} = \begin{pmatrix} \sigma_1 & \omega_1 & & 0 & 0 \\ -\omega_1 & \sigma_1 & & 0 & 0 \\ 0 & 0 & \ddots & \vdots & \vdots \\ 0 & 0 & & \sigma_m & \omega_m \\ 0 & 0 & & -\omega_m & \sigma_m \end{pmatrix} x.$$

In this case, the eigenvalues can be shown to be $\lambda_j = \sigma_j \pm i\omega_j$. We once again can separate the state trajectories into independent solutions for each pair of states, and the solutions are of the form

$$\begin{aligned} x_{2j-1}(t) &= e^{\sigma_j t}(x_{2j-1}(0)\cos\omega_j t + x_{2j}(0)\sin\omega_j t), \\ x_{2j}(t) &= e^{\sigma_j t}(-x_{2j-1}(0)\sin\omega_j t + x_{2j}(0)\cos\omega_j t), \end{aligned}$$

where $j = 1, 2, \ldots, m$. We see that this system is asymptotically stable if and only if $\sigma_j = \operatorname{Re}\lambda_j < 0$. It is also possible to combine real and complex eigenvalues in (block) diagonal form, resulting in a mixture of solutions of the two types.

Very few systems are in one of the diagonal forms above, but some systems can be transformed into these forms via coordinate transformations. One such class of systems is those for which the dynamics matrix has distinct (non-repeating) eigenvalues. In this case there is a matrix $T \in \mathbb{R}^{n\times n}$ such that the matrix TAT^{-1} is in (block) diagonal form, with the block diagonal elements corresponding to the eigenvalues of the original matrix A. If we choose new coordinates $z = Tx$, then

$$\frac{dz}{dt} = T\dot{x} = TAx = TAT^{-1}z$$

and the linear system has a (block) diagonal dynamics matrix. Furthermore, the eigenvalues of the transformed system are the same as the original system since if v is an eigenvector of A, then $w = Tv$ can be shown to be an eigenvector of TAT^{-1}. We can reason about the stability of the original system by noting that $x(t) = T^{-1}z(t)$, and so if the transformed system is stable (or asymptotically stable), then the original system has the same type of stability.

This analysis shows that for linear systems with distinct eigenvalues, the stability of the system can be completely determined by examining the real part of the eigenvalues of the dynamics matrix. For more general systems, we make use of the following theorem, proved in the next chapter:

Theorem 3.1 (Stability of a linear system)**.** *The system*

$$\frac{dx}{dt} = Ax$$

is asymptotically stable if and only if all eigenvalues of A all have a strictly negative real part and is unstable if any eigenvalue of A has a strictly positive real part.

In the case in which the system state is two-dimensional, that is, $x \in \mathbb{R}^2$, we have a simple way of determining the eigenvalues of a matrix A. Specifically, denote by $\mathrm{tr}(A)$ the trace of A, that is, the sum of the diagonal terms, and let $\det(A)$ be the determinant of A. Then, we have that the two eigenvalues are given by

$$\lambda_{1,2} = \frac{1}{2}\left(\mathrm{tr}(A) \pm \sqrt{\mathrm{tr}(A)^2 - 4\det(A)}\right).$$

Both eigenvalues have negative real parts when (1) $\mathrm{tr}(A) < 0$ and (2) $\det(A) > 0$. By contrast, if condition (2) is satisfied but $\mathrm{tr}(A) > 0$, the eigenvalues have positive real parts.

An important feature of differential equations is that it is often possible to determine the local stability of an equilibrium point by approximating the system by a linear system. Suppose that we have a nonlinear system

$$\frac{dx}{dt} = f(x)$$

that has an equilibrium point at x_e. Computing the Taylor series expansion of the vector field, we can write

$$\frac{dx}{dt} = f(x_e) + \left.\frac{\partial f}{\partial x}\right|_{x_e} (x - x_e) + \text{higher-order terms in } (x - x_e).$$

Since $f(x_e) = 0$, we can approximate the system by choosing a new state variable $z = x - x_e$ and writing

$$\frac{dz}{dt} = Az, \qquad \text{where} \quad A = \left.\frac{\partial f}{\partial x}\right|_{x_e}. \tag{3.6}$$

We call the system (3.6) the *linear approximation* of the original nonlinear system or the *linearization* at x_e. We also refer to matrix A as the *Jacobian matrix* of the original nonlinear system.

The fact that a linear model can be used to study the behavior of a nonlinear system near an equilibrium point is a powerful one. Indeed, we can take this even further and use a local linear approximation of a nonlinear system to design a feedback law that keeps the system near its equilibrium point (design of dynamics). Thus, feedback can be used to make sure that solutions remain close to the equilibrium point, which in turn ensures that the linear approximation used to stabilize it is valid.

Example 3.4 (Negative autoregulation)**.** Consider again the negatively autoregulated gene modeled by the equations

$$\frac{dx_1}{dt} = \frac{1}{1+x_2} - x_1, \qquad \frac{dx_2}{dt} = x_1 - x_2.$$

In this case,

$$f(x) = \begin{pmatrix} \frac{1}{1+x_2} - x_1 \\ x_1 - x_2 \end{pmatrix},$$

so that, letting $x_e = (x_{1,e}, x_{2,e})$, the Jacobian matrix is given by

$$A = \left.\frac{\partial f}{\partial x}\right|_{x_e} = \begin{pmatrix} -1 & -\frac{1}{(1+x_{2,e})^2} \\ 1 & -1 \end{pmatrix}.$$

In this case, we have that $\mathrm{tr}(A) = -2 < 0$ and that $\det(A) = 1 + \frac{1}{(1+x_{2,e})^2} > 0$. Hence, independently of the value of the equilibrium point, the eigenvalues have both negative real parts, which implies that the equilibrium point x_e is asymptotically stable. ∇

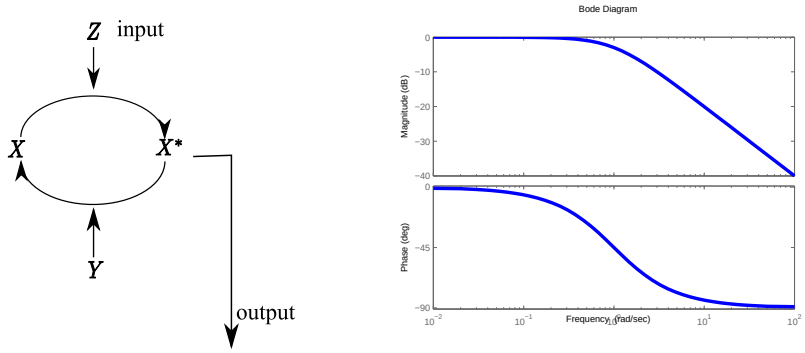
Frequency domain analysis

Frequency domain analysis is a way to understand how well a system can respond to rapidly changing input stimuli. As a general rule, most physical systems display an increased difficulty in responding to input stimuli as the frequency of variation increases: when the input stimulus changes faster than the natural time scales of the system, the system becomes incapable of responding. If instead the input stimulus is changing much slower than the natural time scales of the system, the system will respond very accurately. That is, the system behaves like a "low-pass filter". The cut-off frequency at which the system does not display a significant response is called the *bandwidth* and quantifies the dominant time scale. To identify this dominant time scale, we can perform input/output experiments in which the system is excited with periodic input at various frequencies.

Example 3.5 (Phosphorylation cycle)**.** To illustrate the basic ideas, we consider the frequency response of a phosphorylation cycle, in which enzymatic reactions are modeled by a first order reaction. Referring to Figure 3.6a, we have that the one step reactions involved are given by

$$\mathrm{Z} + \mathrm{X} \xrightarrow{k_1} \mathrm{Z} + \mathrm{X}^*, \qquad \mathrm{Y} + \mathrm{X}^* \xrightarrow{k_2} \mathrm{Y} + \mathrm{X},$$

with conservation law $X + X^* = X_{\mathrm{tot}}$. Let Y_{tot} be the total amount of phosphatase. We assume that the kinase Z has a time-varying concentration, which we view as the *input* to the system, while X^* is the *output* of the system.



(a) Reaction daigram (b) Frequency response

Figure 3.6: (a) Diagram of a phosphorylation cycle, in which Z is the kinase, X is the substrate, and Y is the phosphatase. (b) Bode plot showing the magnitude and phase lag for the frequency response of a one step reaction model of the phosphorylation system on the left. The magnitude is plotted in decibels (dB), in which $M|_{dB} = 20\log_{10}(M)$. The parameters are $\beta = \gamma = 1$.

The differential equation model is given by

$$\frac{dX^*}{dt} = k_1 Z(t)(X_{\text{tot}} - X^*) - k_2 Y_{\text{tot}} X^*,$$

If we assume that the cycle is weakly activated ($X^* \ll X_{\text{tot}}$), the above equation is well approximated by

$$\frac{dX^*}{dt} = \beta Z(t) - \gamma X^*, \tag{3.7}$$

where $\beta = k_1 X_{\text{tot}}$ and $\gamma = k_2 Y_{\text{tot}}$. To determine the frequency response, we set the input $Z(t)$ to a periodic function. It is customary to take sinusoidal functions as the input signal as they lead to an easy way to calculate the frequency response. Let then $Z(t) = A_0 \sin(\omega t)$.

Since equation (3.7) is linear in the state X^* and input Z, it can be directly integrated to lead to

$$X^*(t) = \frac{A_0\beta}{\sqrt{\omega^2+\gamma^2}} \sin(\omega t - \tan^{-1}(\omega/\gamma)) - \frac{A_0\beta\omega}{(\omega^2+\gamma^2)} e^{-\gamma t}.$$

The second term dies out for t large enough. Hence, the steady state response is given by the first term. The amplitude of response is thus given by $A_0\beta/\sqrt{\omega^2+\gamma^2}$, in which the gain $\beta/\sqrt{\omega^2+\gamma^2}$ depends on the system parameters and on the frequency of the input stimulation.

As this frequency increases, the amplitude decreases and approaches zero for infinite frequencies. Also, the argument of the sine function shows a negative phase shift of $\tan^{-1}(\omega/\gamma)$, which indicates that there is an increased delay in responding

to the input as the frequency increases. Hence, the key quantities in the frequency response are the magnitude $M(\omega)$, also called gain of the system, and phase lag $\phi(\omega)$ given by

$$M(\omega) = \frac{\beta}{\sqrt{\omega^2 + \gamma^2}}, \qquad \phi(\omega) = \tan^{-1}\left(\frac{\omega}{\gamma}\right).$$

These are plotted in Figure 3.6b, a type of figure known as a *Bode plot*.

The bandwidth of the system, denoted ω_B is the frequency at which the gain drops below $M(0)/\sqrt{2}$. In this case, the bandwidth is given by $\omega_B = \gamma = k_2 Y_{\text{tot}}$, which implies that the bandwidth of the system can be made larger by increasing the amount of phosphatase. However, note that since $M(0) = \beta/\gamma = k_1 X_{\text{tot}}/(k_2 Y_{\text{tot}})$, increased phosphatase will also result in decreased amplitude of response. Hence, if one wants to increase the bandwidth of the system while keeping the value of $M(0)$ (also called the *zero frequency gain*) unchanged, one should increase the total amounts of substrate and phosphatase in comparable proportions. Fixing the value of the zero frequency gain, the bandwidth of the system increases with increased amounts of phosphatase and kinase. ∇

More generally, the *frequency response* of a linear system with one input and one output

$$\dot{x} = Ax + Bu, \qquad y = Cx + Du$$

is the response of the system to a sinusoidal input $u = a\sin\omega t$ with input amplitude a and frequency ω. The *transfer function* for a linear system is given by

$$G_{yu}(s) = C(sI - A)^{-1}B + D$$

and represents the response of a system to an exponential signal of the form $u(t) = e^{st}$ where $s \in \mathbb{C}$. In particular, the response to a sinusoid $u = a\sin\omega t$ is given by $y = Ma\sin(\omega t + \phi)$ where the gain M and phase shift ϕ can be determined from the transfer function evaluated at $s = i\omega$:

$$G_{yu}(i\omega) = Me^{i\phi}, \qquad \begin{aligned} M(\omega) &= |G_{yu}(i\omega)| = \sqrt{\text{Im}(G_{yu}(i\omega))^2 + \text{Re}(G_{yu}(i\omega))^2} \\ \phi(\omega) &= \tan^{-1}\left(\frac{\text{Im}(G_{yu}(i\omega))}{\text{Re}(G_{yu}(i\omega))}\right), \end{aligned}$$

where $\text{Re}(\cdot)$ and $\text{Im}(\cdot)$ represent the real and imaginary parts of a complex number.

For finite dimensional linear (or linearized) systems, the transfer function be written as a ratio of polynomials in s:

$$G(s) = \frac{b(s)}{a(s)}.$$

The values of s at which the numerator vanishes are called the *zeros* of the transfer function and the values of s at which the denominator vanishes are called the *poles*.

The transfer function representation of an input/output linear system is essentially equivalent to the state space description, but we reason about the dynamics by looking at the transfer function instead of the state space matrices. For example, it can be shown that the poles of a transfer function correspond to the eigenvalues of the matrix A, and hence the poles determine the stability of the system. In addition, interconnections between subsystems often have simple representations in terms of transfer functions. For example, two systems G_1 and G_2 in series (with the output of the first connected to the input of the second) have a combined transfer function $G_{\text{series}}(s) = G_1(s)G_2(s)$ and two systems in parallel (a single input goes to both systems and the outputs are summed) has the transfer function $G_{\text{parallel}}(s) = G_1(s) + G_2(s)$.

Transfer functions are useful representations of linear systems because the properties of the transfer function can be related to the properties of the dynamics. In particular, the shape of the frequency response describes how the system response to inputs and disturbances, as well as allows us to reason about the stability of interconnected systems. The Bode plot of a transfer function gives the magnitude and phase of the frequency response as a function of frequency and the *Nyquist plot* can be used to reason about stability of a closed loop system from the open loop frequency response (AM08, Section 9.2).

Returning to our analysis of biomolecular systems, suppose we have a systems whose dynamics can be written as

$$\dot{x} = f(x, \theta, u)$$

and we wish to understand how the solutions of the system depend on the parameters θ and input disturbances u. We focus on the case of an equilibrium solution $x(t; x_0, \theta_0) = x_e$. Let $z = x - x_e$, $\tilde{u} = u - u_0$ and $\tilde{\theta} = \theta - \theta_0$ represent the deviation of the state, input and parameters from their nominal values. Linearization can be performed in a way similar to the way it was performed for a system with no inputs. Specifically, we can write the dynamics of the perturbed system using its linearization as

$$\frac{dz}{dt} = \left(\frac{\partial f}{\partial x}\right)_{(x_e,\theta_0,u_0)} \cdot z \quad + \quad \left(\frac{\partial f}{\partial \theta}\right)_{(x_e,\theta_0,u_0)} \cdot \tilde{\theta} \quad + \quad \left(\frac{\partial f}{\partial w}\right)_{(x_e,\theta_0,u_0)} \cdot \tilde{u}.$$

This linear system describes small deviations from $x_e(\theta_0, w_0)$ but allows $\tilde{\theta}$ and $\tilde{w}$ to be time-varying instead of the constant case considered earlier.

To analyze the resulting deviations, it is convenient to look at the system in the frequency domain. Let $y = Cx$ be a set of values of interest. The transfer functions between $\tilde{\theta}$, $\tilde{w}$ and y are given by

$$H_{y\tilde{\theta}}(s) = C(sI - A)^{-1}B_\theta, \qquad H_{y\tilde{w}}(s) = C(sI - A)^{-1}B_w,$$

where

$$A = \left.\frac{\partial f}{\partial x}\right|_{(x_e,\theta_0,w_0)}, \qquad B_\theta = \left.\frac{\partial f}{\partial \theta}\right|_{(x_e,\theta_0,w_0)}, \qquad B_w = \left.\frac{\partial f}{\partial w}\right|_{(x_e,\theta_0,w_0)}.$$

Note that if we let $s = 0$, we get the response to small, constant changes in parameters. For example, the change in the outputs y as a function of constant changes in the parameters is given by

$$H_{y\tilde{\theta}}(0) = CA^{-1}B_\theta = CS_{x,\theta}.$$

Example 3.6 (Transcriptional regulation)**.** Consider a genetic circuit consisting of a single gene. The dynamics of the system are given by

$$\frac{dm_{\mathrm{P}}}{dt} = F(P) - \delta m_{\mathrm{P}}, \qquad \frac{dP}{dt} = \kappa m_{\mathrm{P}} - \gamma P,$$

where m_{P} is the mRNA concentration and P is the protein concentration. Suppose that the mRNA degradation rate δ can change as a function of time and that we wish to understand the sensitivity with respect to this (time-varying) parameter. Linearizing the dynamics around an equilibrium point

$$A = \begin{pmatrix} -\delta & F'(P_e) \\ \kappa & -\gamma \end{pmatrix}, \qquad B_\delta = \begin{pmatrix} -m_{\mathrm{P},e} \\ 0 \end{pmatrix}.$$

For the case of no feedback we have $F(P) = \alpha$, and the system has an equilibrium point at $m_{\mathrm{P},e} = \alpha/\delta$, $P_e = \kappa\alpha/(\gamma\delta)$. The transfer function from δ to P, after linearization about the steady state, is given by

$$G^{\mathrm{ol}}_{P\delta}(s) = \frac{-\kappa m_{\mathrm{P},e}}{(s+\delta)(s+\gamma)},$$

where δ_0 represents the nominal value of δ around which we are linearizing. For the case of negative regulation, we have

$$F(P) = \frac{\alpha}{1+(P/K)^n} + \alpha_0,$$

and the resulting transfer function is given by

$$G^{\mathrm{cl}}_{P\delta}(s) = \frac{\beta m_{\mathrm{P},e}}{(s+\delta_0)(s+\gamma)+\kappa\sigma}, \qquad \sigma = -F'(P_e) = \frac{n\alpha P_e^{n-1}/K^n}{(1+P_e^n/K^n)^2}.$$

Figure 3.7 shows the frequency response for the two circuits. We see that the feedback circuit attenuates the response of the system to perturbations with low-frequency content but slightly amplifies perturbations at high frequency (compared to the open loop system). ∇

3.2 Robustness

The term "robustness" refers to the general ability of a system to continue to function in the presence of uncertainty. In the context of this text, we will want to be

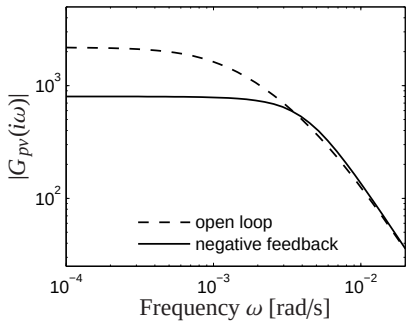


Figure 3.7: Attenuation of perturbations in a genetic circuit.

more precise. We say that a given function (of the circuit) is robust with respect to a set of specified perturbations if the sensitivity of that function to perturbations is small. Thus, to study robustness, we must specify both the function we are interested in and the set of perturbations that we wish to consider.

In this section we study the robustness of the system

$$\dot{x} = f(x, \theta, u), \qquad y = h(x, \theta)$$

to various perturbations in the parameters θ and disturbance inputs u. The function we are interested in is modeled by the outputs y and hence we seek to understand how y changes if the parameters θ are changed by a small amount or if external disturbances u are present. We say that a system is robust with respect to these perturbations if y undergoes little changes as these perturbations are introduced.

Parametric uncertainty

In addition to studying the input/output transfer curve and the stability of a given equilibrium point, we can also study how these features change with respect to changes in the system parameters θ. Let $y_e(\theta_0, u_0)$ represent the output corresponding to an equilibrium point x_e with fixed parameters θ_0 and external input u_0, so that $f(x_e, \theta_0, u_0) = 0$. We assume that the equilibrium point is stable and focus here on understanding how the value of the output, the location of the equilibrium point and the dynamics near the equilibrium point vary as a function of changes in the parameters θ and external inputs w.

We start by assuming that $u = 0$ and investigating how x_e and y_e depend on θ. The simplest approach is to analytically solve the equation $f(x_e, \theta_0) = 0$ for x_e and then set $y_e = h(x_e, \theta_0)$. However, this is often difficult to do in closed form and so as an alternative we instead look at the linearized response given by

$$S_{x,\theta} := \left.\frac{dx_e}{d\theta}\right|_{\theta_0}, \qquad S_{y,\theta} := \left.\frac{dy_e}{d\theta_0}\right|_{\theta_0},$$

which is the (infinitesimal) change in the equilibrium state and the output due to a change in the parameter. To determine $S_{x,\theta}$ we begin by differentiating the relationship $f(x_e(\theta),\theta) = 0$ with respect to θ:

$$\frac{df}{d\theta} = \frac{\partial f}{\partial x}\frac{dx_e}{d\theta} + \frac{\partial f}{\partial \theta} = 0 \quad \Longrightarrow \quad S_{x,\theta} = \frac{dx_e}{d\theta} = -\left(\frac{\partial f}{\partial x}\right)^{-1}\frac{\partial f}{\partial \theta}\bigg|_{(xe,\theta_0)}. \tag{3.8}$$

Similarly, we can compute the change in the output sensitivity as

$$S_{y,\theta} = \frac{dy_e}{d\theta} = \frac{\partial h}{\partial x}\frac{dx_e}{d\theta} + \frac{\partial h}{\partial \theta} = -\left(\frac{\partial h}{\partial x}\left(\frac{\partial f}{\partial x}\right)^{-1}\frac{\partial f}{\partial \theta} + \frac{\partial h}{\partial \theta}\right)\bigg|_{(xe,\theta_0)}.$$

These quantities can be computed numerically and hence we can evaluate the effect of small (but constant) changes in the parameters θ on the equilibrium state x_e and corresponding output value y_e.

A similar analysis can be performed to determine the effects of small (but constant) changes in the external input u. Suppose that x_e depends on both θ and u, with $f(x_e,\theta_0,u_0) = 0$ and θ_0 and u_0 representing the nominal values. Then

$$\frac{dx_e}{d\theta} = -\left(\frac{\partial f}{\partial x}\right)^{-1}\frac{\partial f}{\partial \theta}\bigg|_{(xe,\theta_0,u_0)}, \qquad \frac{dx_e}{du} = -\left(\frac{\partial f}{\partial x}\right)^{-1}\frac{\partial f}{\partial u}\bigg|_{(xe,\theta_0,u_0)}.$$

The sensitivity matrix can be normalized by dividing the parameters by their nominal values and rescaling the outputs (or states) by their equilibrium values. If we define the scaling matrices

$$D^{x_e} = \mathrm{diag}\{x_e\}, \qquad D^{y_e} = \mathrm{diag}\{y_e\}, \qquad D^{\theta} = \mathrm{diag}\{\theta\},$$

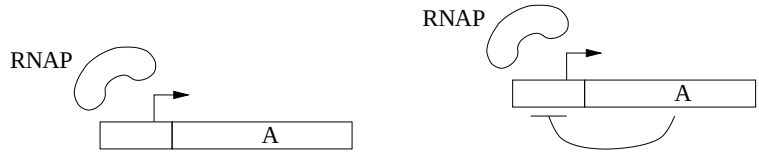
Then the scaled sensitivity matrices can be written as

$$\bar{S}_{x,\theta} = (D^{x_e})^{-1} S_{x,\theta} D^{\theta}, \qquad \bar{S}_{y,\theta} = (D^{y_e})^{-1} S_{y,\theta} D^{\theta}. \tag{3.9}$$

The entries in this matrix describe how a fractional change in a parameter gives a fractional change in the output, relative to the nominal values of the parameters and outputs.

Example 3.7 (Transcriptional regulation)**.** Consider again the case of transcriptional regulation described in Example 3.6. We wish to study the response of the protein concentration to fluctuations in its parameters in two cases: a *constitutive promoter* (no regulation) and self-repression (negative feedback), illustrated in Figure 3.8. For the case of no feedback we have $F(p) = \alpha$, and the system has an equilibrium point at $m_e = \alpha/\delta$, $P_e = \kappa\alpha/(\gamma\delta)$. The parameter vector can be taken as $\theta = (\alpha,\delta,\kappa,\gamma)$. Since we have a simple expression for the equilibrium concentrations, we can compute the sensitivity to the parameters directly:

$$\frac{\partial x_e}{\partial \theta} = \begin{pmatrix} \frac{1}{\delta} & -\frac{\alpha}{\delta^2} & 0 & 0 \\ \frac{\kappa}{\gamma\delta} & -\frac{\beta\alpha}{\gamma\delta^2} & \frac{\alpha}{\gamma\delta} & -\frac{\kappa\alpha}{\delta\gamma^2} \end{pmatrix},$$



(a) Open loop (b) Negative feedback

Figure 3.8: Parameter sensitivity in a genetic circuit. The open loop system (a) consists of a constitutive promoter, while the closed loop circuit (b) is self-regulated with negative feedback (repressor).

where the parameters are evaluated at their nominal values, but we leave off the subscript 0 on the individual parameters for simplicity. If we choose the parameters as $\theta_0 = (0.00138, 0.00578, 0.115, 0.00116)$, then the resulting sensitivity matrix evaluates to

$$S^{\text{open}}_{x_e,\theta} \approx \begin{pmatrix} 170 & -41 & 0 & 0 \\ 17000 & -4100 & 210 & -21000 \end{pmatrix}. \tag{3.10}$$

If we look instead at the scaled sensitivity matrix, then the open loop nature of the system yields a particularly simple form:

$$\bar{S}^{\text{open}}_{x_e,\theta} = \begin{pmatrix} 1 & -1 & 0 & 0 \\ 1 & -1 & 1 & -1 \end{pmatrix}. \tag{3.11}$$

In other words, a 10% change in any of the parameters will lead to a comparable positive or negative change in the equilibrium values.

For the case of negative regulation, we have

$$F(P) = \frac{\alpha}{1 + (P/K)^n} + \alpha_0,$$

and the equilibrium points satisfy

$$m_e = \frac{\gamma}{\beta} P_e, \qquad \frac{\alpha}{1 + P_e^n/K^n} + \alpha_0 = \delta m_e = \frac{\delta\gamma}{\beta} P_e. \tag{3.12}$$

In order to make a proper comparison with the previous case, we need to be careful to choose the parameters so that the equilibrium concentration P_e matches that of the open loop system. We can do this by modifying the promoter strength α or the RBS strength κ so that the second formula in equation (3.12) is satisfied or, equivalently, choose the parameters for the open loop case so that they match the closed loop steady state protein concentration (see Example 2.2).

Rather than attempt to solve for the equilibrium point in closed form, we instead investigate the sensitivity using the computations in equation (3.12). The state, dynamics and parameters are given by

$$x = \begin{pmatrix} m & P \end{pmatrix}, \qquad f(x,\theta) = \begin{pmatrix} F(P) - \delta m \\ \kappa m - \gamma P \end{pmatrix}, \qquad \theta = \begin{pmatrix} \alpha_0 & \delta & \kappa & \gamma & \alpha & n & K \end{pmatrix}.$$

Note that the parameters are ordered such that the first four parameters match the open loop system. The linearizations are given by

$$\frac{\partial f}{\partial x} = \begin{pmatrix} -\delta & F'(P_e) \\ \beta & -\gamma \end{pmatrix}, \qquad \frac{\partial f}{\partial \theta} = \begin{pmatrix} 1 & -m & 0 & 0 & \frac{1}{1+(P/K)^n} & \frac{K^n \alpha P^n \log(P)}{(K^n+P^n)^2} & \frac{\alpha}{(1+(P/K)^n)^2} \\ 0 & 0 & m & -P & 0 & 0 & 0 \end{pmatrix},$$

where again the parameters are taken to be their nominal values. From this we can compute the sensitivity matrix as

$$S_{x,\theta} = \begin{pmatrix} -\frac{\gamma}{\gamma\delta-\kappa F'} & \frac{\gamma m}{\gamma\delta-\kappa F'} & -\frac{mF'}{\gamma\delta-\kappa F'} & \frac{PF'}{\gamma\delta-\kappa F'} & -\frac{\gamma\frac{\partial F}{\partial \alpha_1}}{\gamma\delta-\kappa F'} & -\frac{\gamma\frac{\partial F}{\partial n}}{\gamma\delta-\kappa F'} & -\frac{\gamma\frac{\partial F}{\partial K}}{\gamma\delta-\kappa F'} \\ -\frac{\kappa}{\gamma\delta-\kappa F'} & \frac{\kappa m}{\gamma\delta-\kappa F'} & -\frac{\delta m}{\gamma\delta-\kappa F'} & \frac{\delta P}{\gamma\delta-\kappa F'} & -\frac{\kappa\frac{\partial F}{\partial \alpha_1}}{\gamma\delta-\kappa F'} & -\frac{\kappa\frac{\partial F}{\partial n}}{\gamma\delta-\kappa F'} & -\frac{\kappa\frac{\partial F}{\partial K}}{\gamma\delta-\kappa F'} \end{pmatrix},$$

where $F' = \partial F/\partial P$ and all other derivatives of F are evaluated at the nominal parameter values.

We can now evaluate the sensitivity at the same protein concentration as we use in the open loop case. The equilibrium point is given by

$$x_e = \begin{pmatrix} m_e \\ P_e \end{pmatrix} = \begin{pmatrix} \frac{\alpha}{\delta} \\ \frac{\alpha\kappa}{\gamma\delta} \end{pmatrix} = \begin{pmatrix} 0.239 \\ 23.9 \end{pmatrix}$$

and the sensitivity matrix is

$$S^{\text{closed}}_{x_e,\theta} \approx \begin{pmatrix} 76.1 & -18.2 & -1.16 & 116. & 0.134 & -0.212 & -0.000117 \\ 7610. & -1820. & 90.8 & -9080. & 13.4 & -21.2 & -0.0117 \end{pmatrix}.$$

The scaled sensitivity matrix becomes

$$\bar{S}^{\text{closed}}_{x_e,\theta} \approx \begin{pmatrix} 0.16 & -0.44 & -0.56 & 0.56 & 0.28 & -1.78 & -3.08\times10^{-7} \\ 0.16 & -0.44 & 0.44 & -0.44 & 0.28 & -1.78 & -3.08\times10^{-7} \end{pmatrix}. \quad (3.13)$$

Comparing this equation with equation (3.11), we see that there is reduction in the sensitivity with respect to most parameters. In particular, we become less sensitive to those parameters that are not part of the feedback (columns 2–4), but there is higher sensitivity with respect to some of the parameters that are part of the feedback mechanisms (particularly n). ∇

More generally, we may wish to evaluate the sensitivity of a (non-constant) solution to parameter changes. This can be done by computing the function $dx(t)/d\theta$, which describes how the state changes at each instant in time as a function of (small) changes in the parameters θ.

Let $x(t,\theta_0)$ be a solution of the nominal system

$$\dot{x} = f(x,\theta_0,u), \qquad x(0) = x_0.$$

To compute $dx/d\theta$, we write a differential equation for how it evolves in time:

$$\frac{d}{dt}\left(\frac{dx}{d\theta}\right) = \frac{d}{d\theta}\left(\frac{dx}{dt}\right) = \frac{d}{d\theta}\left(f(x,\theta,u)\right) = \frac{\partial f}{\partial x}\frac{dx}{d\theta} + \frac{\partial f}{\partial \theta}.$$

This is a differential equation with $n \times m$ states given by the entries of the matrix $S_{x,\theta}(t) = dx_i(t)/d\theta_j$ and with initial condition $S_{x,\theta}(0) = 0$ (since changes to the parameters to not affect the initial conditions).

To solve these equations, we must simultaneously solve for the state x and the sensitivity $S_{x,\theta}$ (whose dynamics depend on x). Thus, letting

$$M(t,\theta_0) := \frac{\partial f}{\partial x}(x,\theta,u)\bigg|_{x=x(t,\theta_0),\theta=\theta_0}, \qquad N(t,\theta_0) := \frac{\partial f}{\partial \theta}(x,\theta,u)\bigg|_{x=x(t,\theta_0),\theta=\theta_0},$$

we solve the set of $n + nm$ coupled differential equations

$$\frac{dx}{dt} = f(x,\theta_0,u), \qquad \frac{dS_{x,\theta}}{dt} = M(t,\theta_0)S_{x,\theta} + N(t,\theta_0), \tag{3.14}$$

with initial condition $x(0) = x_0$ and $S_{x,\theta}(0) = 0$.

This differential equation generalizes our previous results by allowing us to evaluate the sensitivity around a (non-constant) trajectory. Note that in the special case that we are at an equilibrium point and the dynamics for $S_{x,\theta}$ are stable, the steady state solution of equation (3.14) is identical to that obtained in equation (3.8). However, equation (3.14) is much more general, allowing us to determine the change in the state of the system at a fixed time T, for example. This equation also does not require that our solution stay near an equilibrium point, it only requires that our perturbations in the parameters are sufficiently small.

Several simulation tools include the ability to do sensitivity analysis of this sort, including `COPASI`.

Adaptation and disturbance rejection

A system is said to adapt to the input u when the steady state value of its output y is independent of the actual (constant) non-zero value of the input (Figure 3.9). Basically, after the input changes to a constant non-zero value, the output returns to its original value after a transient perturbation. Adaptation corresponds to the concept of *disturbance rejection* in control theory. The full notion of disturbance rejection is more general and depends on the specific disturbance input and it is studied using the internal model principle [83].

For example, for adaptation to constant signals, the internal model principle requires integral feedback. The internal model principle is a powerful way to uncover biochemical structures in natural networks that are known to have the adaptation property. An example of this is the bacterial chemotaxis described in more detail in Chapter 5.

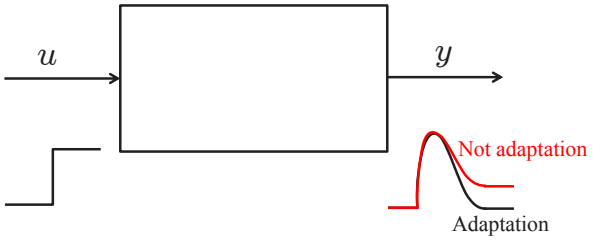


Figure 3.9: Adaptation property. The system is said to have the adaptation property if the steady state value of the output does not depend on the steady state value of the input. Hence, after a constant input perturbation, the output returns to its original value.

We illustrate two main mechanisms to attain adaptation: integral feedback and incoherent feedforward loops (IFFLs). We next study these two mechanisms from a mathematical standpoint to illustrate how they achieve adaptation. Possible biomolecular implementations are presented in later chapters.

Integral feedback

In integral feedback systems, a "memory" variable z keeps track of the accumulated difference between $y(t)$ and its nominal steady state value y_0. A comparison is performed between this memory variable and the current input u, providing an error term that is used to drive the feedback mechanism that brings the system output back to the desired value y_0 (Figure 3.10).

In this system, the output $y(t)$, after any constant input perturbation u, tends to

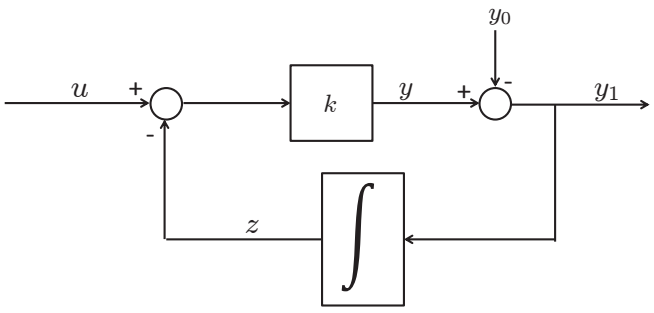


Figure 3.10: Basic block diagram representing a system with integral action.

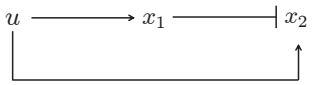


Figure 3.11: Incoherent feedforward loop. The input u affects the output through two channels. It indirectly represses it through an intermediate variable x_1 and it activates it directly.

y_0 for $t \to \infty$ independently of the (constant) value of u. The equations representing the system are given by:

$$\frac{dz}{dt} = y_1, \qquad y_1 = y - y_0, \qquad y = k(u - z),$$

so that the equilibrium is obtained by setting $\dot{z} = 0$, from which we obtain $y = y_0$. That is, the steady state of y does not depend on u. The additional question to answer is whether, after a perturbation u occurs, $y_1(t)$ tends to zero for $t \to \infty$. This is the case if and only if $\dot{z} \to 0$ as $t \to \infty$, which is satisfied if the equilibrium of the system $\dot{z} = -kz + ku - y_0$ is asymptotically stable. This, in turn, is satisfied whenever $k > 0$ and u is a constant. Hence, after a constant perturbation u is applied, the system output y approaches back its original steady state value y_0, that is, y is robust to constant perturbations.

More generally, a system with integral action can take the form

$$\frac{dx}{dt} = f(x,u,k), \quad y = h(x), \quad \frac{dz}{dt} = y - y_0, \quad k = k(x,z),$$

in which the steady state value of y, being the solution to $y - y_0 = 0$, does not depend on u. In turn, y tends to this steady state value for $t \to \infty$ if and only if $\dot{z} \to 0$ as $t \to \infty$. This, in turn, is the case if z tends to a constant value for $t \to \infty$, which is satisfied if u is a constant and the steady state of the above system is asymptotically stable.

Integral feedback is recognized as a key mechanism of perfectly adapting biological systems, both at the physiological level and at the cellular level, such as in blood calcium homeostasis [21], in the regulation of tryptophan in *E. coli* [90], in neuronal control of the prefrontal cortex [65], and in *E. coli* chemotaxis [98].

Incoherent feedforward loops

Feedforward motifs (Figure 3.11) are common in transcriptional networks and it has been shown they are over-represented in *E. coli* gene transcription networks, compared to other motifs composed of three nodes [3]. These are systems in which the input u directly helps promote the production of the output x_2 and also acts as a

delayed inhibitor of the output through an intermediate variable x_1. This incoherent counterbalance between positive and negative effects gives rise, under appropriate conditions, to adaptation. A large number of incoherent feedforward loops participate in important biological processes such as the EGF to ERK activation [70], the glucose to insulin release [71], ATP to intracellular calcium release [61], microRNA regulation [89], and many others.

Several variants of incoherent feedforward loops exist for perfect adaptation. The "sniffer", for example, is one in which the intermediate variable promotes degradation:

$$\frac{dx_1}{dt} = \alpha u - \gamma x_1, \qquad \frac{dx_2}{dt} = \beta u - \delta x_1 x_2. \tag{3.15}$$

In this system, the steady state value of the output x_2 is obtained by setting the time derivatives to zero. Specifically, we have that $\dot{x}_1 = 0$ given $x_1 = \alpha u/\gamma$ and $\dot{x}_2 = 0$ gives $x_2 = \beta u/(\delta x_1)$. IN the case in which $u \neq 0$, these can be combined to yield $x_2 = (\beta\gamma)/(\delta\alpha)$, which is a constant independent of the input u. The linearization of the system at the equilibrium is given by

$$A = \begin{pmatrix} -\gamma & 0 \\ -\delta(\beta\gamma)/(\delta\alpha) & -\delta(\alpha u/\gamma) \end{pmatrix},$$

which has eigenvalues $-\gamma$ and $-\delta(\alpha u/\gamma)$. Since these are both negative, the equilibrium point is asymptotically stable. The sniffer appears in models of neutrophil motion and *Dictyostelium* chemotaxis [97]. Note that in the case in which, for example, u goes back to zero after a perturbation, as it is in the case of a pulse, the output x_2 does not reach back necessarily its original steady state. That is, this system "adapts" only to constant non-zero input stimuli but is not capable of adapting to pulses. This can be seen from equation (3.15), which, when $u = 0$ admits multiple steady states. For more details on this "memory" effect, the reader is referred to [85].

Another form for a feedforward loop is one in which the intermediate variable x_1 inhibits production of the output x_2, such as in the system:

$$\frac{dx_1}{dt} = \alpha u - \gamma x_1, \qquad \frac{dx_2}{dt} = \beta \frac{u}{x_1} - \delta x_2. \tag{3.16}$$

The equilibrium point of this system for a constant non-zero input u is given by setting the time derivatives to zero. From $\dot{x}_1 = 0$, one obtains $x_1 = \alpha u/\gamma$ and from $\dot{x}_2 = 0$ one obtains that $x_2 = \beta u/(\delta x_1)$, which combined together result in $x_2 = (\beta\gamma)/(\delta\alpha)$, which is a constant independent of the input u.

By calculating the linearization at the equilibrium, one obtains

$$A = \begin{pmatrix} -\gamma & 0 \\ -\frac{u}{x_1^2} & -\delta \end{pmatrix},$$

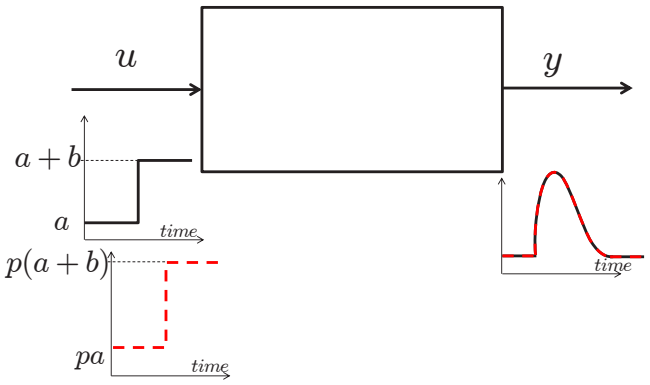


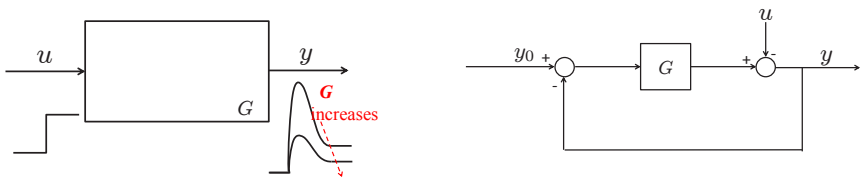
Figure 3.12: Fold-change detection. The output response does not depend on the absolute magnitude of the input but only on the fold change of the input.

whose eigenvalues are given by $-\gamma$ and $-\delta$. Hence, the equilibrium point is asymptotically stable. Further, one can show that the equilibrium point is globally asymptotically stable because the x_1 subsystem is linear, stable, and x_1 approaches a constant value (for constant u) and the x_2 subsystem, in which $\beta u/x_1$ is viewed as an external input is also linear and exponentially stable.

Scale Invariance and fold-change detection

Scale invariance is the property by which the output $x_2(t)$ of the system does not depend on the amplitude of the input $u(t)$ (Figure 3.12). Specifically, consider an adapting system and assume that it pre-adapted to a constant background value a, then apply input $a+b$ and let $x_2(t)$ be the resulting output. Now consider a new background value $p\,a$ for the input and let the system pre-adapt to it. Then apply the input $p(a+b)$ and let $\bar{x}_2(t)$ be the resulting output. The system has the scale invariance property if $x_2(t) = \bar{x}_2(t)$. This also means that the output responds in the same way to inputs changing by the same multiplicative factor (fold), hence this property is also called fold-change detection. Looking at Figure 3.12, the output would present different pulses for different fold changes b/a.

Some incoherent feedforward loops can implement the fold-change detection property [32]. As an example, consider the feedforward motif represented by equations (3.16) and consider two inputs: $u_1(t) = a + b_1(t - t_0)$ and $u_2(t) = pa + pb_1(t - t_0)$. Assume also, as said above, that at time t_0 the system is at the steady state, that is, it pre-adapted. Hence, we have that the two steady states from which the system starts at $t = t_0$ are given by $x_{1,1} = a\alpha/\gamma$ and $x_{1,2} = pa\alpha/\gamma$ for the x_1 variable and by $x_{2,1} = x_{2,2} = (\beta\gamma)/(\delta\alpha)$ for the x_2 variable. Integrating system (**??**) starting from



(a) (b)

Figure 3.13: (a) Disturbance attenuation. A system is said to have the disturbance attenuation property if there is an internal system parameter G such that the system output response becomes arbitrarily close to a nominal output (independent of the input u) by increasing the value of G. (b) High gain feedback. A possible mechanism to attain disturbance attenuation is to feedback the error between the nominal output y_0 and the actual output y through a large gain G.

these initial conditions, we obtain for $t \geq t_0$

$$x_{1,1}(t) = a\frac{\alpha}{\gamma}e^{-\gamma(t-t_0)} + (a+b)(1-e^{-\gamma(t-t_0)}) \quad \text{and}$$
$$x_{1,2}(t) = pa\frac{\alpha}{\gamma}e^{-\gamma(t-t_0)} + p(a+b)(1-e^{-\gamma(t-t_0)}).$$

Using these in the expression of $\dot{x}_2$ in equation (**??**) gives the differential equations to which $x_{2,1}(t)$ and $x_{2,2}(t)$ obey for $t \geq t_0$ as

$$\frac{dx_{2,1}}{dt} = \frac{\beta(a+b)}{a\frac{\alpha}{\gamma}e^{-\gamma(t-t_0)} + (a+b)(1-e^{-\gamma(t-t_0)})} - \delta x_{2,1}, \qquad x_{2,1}(t_0) = (\beta\gamma)/(\delta\alpha)$$

and

$$\frac{dx_{2,2}}{dt} = \frac{p\beta(a+b)}{pa\frac{\alpha}{\gamma}e^{-\gamma(t-t_0)} + p(a+b)(1-e^{-\gamma(t-t_0)})} - \delta x_{2,2}, \qquad x_{2,2}(t_0) = (\beta\gamma)/(\delta\alpha),$$

which give $x_{2,1}(t) = x_{2,2}(t)$ for all $t \geq t_0$. Hence, the system responds exactly the same way after changes in the input of the same fold. The output response is not dependent on the scale of the input but only on its shape.

Disturbance attenuation

A system has the property of disturbance attenuation if there is a system parameter G such that the output response $y(t)$ to the input $u(t)$ can be made arbitrarily small as G is increased (Figure 3.13a). A possible mechanism for disturbance attenuation is high gain feedback (Figure 3.13b). In a high gain feedback configuration, the error between the output y, perturbed by some exogenous disturbance u, and a desired

nominal output y_0 is fed back with a negative sign to produce the output y itself. If $y_0 > y$, this will result in an increase of y, otherwise it will result in a decrease of y. Mathematically, one obtains from the block diagram that

$$y = \frac{u}{1+G} + y_0 \frac{G}{1+G},$$

so that as G increases the (relative) contribution of u on the output of the system can be arbitrarily reduced.

High gain feedback can take a much more general form. Consider a system with $x \in \mathbb{R}^n$ in the form $\dot{x} = f(x,t)$. We say that this system is *contracting* if any two trajectories starting from different initial conditions tend to each other as time increase to infinity. A sufficient condition for the system to be contracting is that in some set of coordinates, with matrix transformation denoted Θ, the symmetric part of the linearization matrix (Jacobian) is negative definite. That is, that the largest eigenvalue of

$$\frac{1}{2}\left(\frac{\partial f}{\partial x} + \frac{\partial f}{\partial x}^T\right),$$

is negative. We denote this eigenvalue by $-\lambda$ for $\lambda > 0$ and call it the contraction rate of the system.

Now, consider the nominal system $\dot{x} = Gf(x,t)$ for $G > 0$ and its perturbed version $\dot{x}_p = Gf(x_p,t) + u(t)$. Assume that the input $u(t)$ is bounded everywhere in norm by a constant $C > 0$. If the system is contracting, we have the following robustness result:

$$\|x(t) - x_p(t)\| \leq \chi \|x(0) - x_p(0)\| e^{-G\lambda t} + \frac{\chi C}{\lambda G},$$

in which χ is an upper bound on the condition number (ratio between the largest and the smallest eigenvalue of $\Theta^T\Theta$) of the transformation matrix Θ [57]. Hence, if the perturbed and the nominal systems start from the same initial conditions, the difference between their states can be made arbitrarily small by increasing the gain G. Hence, the system has the disturbance attenuation property.

A comprehensive treatment of concepts of stability and robustness can be found in standard references [51, 84].

3.3 Analysis of Reaction Rate Equations

The previous section considered analysis techniques for general dynamical systems with small perturbations. In this section, we specialize to the case where the dynamics have the form of a reaction rate equation:

$$\frac{ds}{dt} = Nv(x,\theta), \tag{3.17}$$

where x is the vector of species concentrations, θ is the vector of reaction parameters, N is the stoichiometry matrix and $v(x,\theta)$ is the reaction rate (or flux) vector.

Reduced reaction dynamics

When analyzing reaction rate equations, it is often the case that there are conserved quantities in the dynamics. For example, conservation of mass will imply that if all compounds containing a given species are captured by the model, the total mass of that species will be constant. This type of constraint will then give a conserved quantity of the form $c_i = H_i x$ where H_i represents that combinations of species in which the given element appears. Since c_i is constant, it follows that $dc_i/dt = 0$ and, aggregating the set of all conserved species, we have

$$0 = \frac{dc}{dt} = H\frac{ds}{dt} = HNv(x,\theta) \quad \text{for all } x.$$

If we assume that the vector of fluxes spans $\mathbb{R}^m$ (the range of $v : \mathbb{R}^n \times \mathbb{R}^p \to \mathbb{R}^m$), then this implies that the conserved quantities correspond to the left null space of the stoichiometry matrix N.

It is often useful to remove the conserved quantities from the description of the dynamics and write the dynamics for a set of independent species. To do this, we transform the state of the system into two sets of variables:

$$\begin{pmatrix} x_i \\ x_d \end{pmatrix} = \begin{pmatrix} P \\ H \end{pmatrix} x. \tag{3.18}$$

The vector $x_i = Px$ is the set of independent species and is typically chosen as a subset of the original species of the model (so that the rows P consists of all zeros and a single 1 in the column corresponding to the selected species). The matrix H should span the left null space of N, so that x_d represents the set of dependent concentrations. These dependent species do not necessarily correspond to individual species, but instead are often combinations of species (for example, the total concentration of a given element that appears in a number of molecules that participate in the reaction).

Given the decomposition (3.18), we can rewrite the dynamics of the system in terms of the independent variables x_i. We start by noting that given x_i and x_d, we can reconstruct the full set of species x:

$$x = \begin{pmatrix} P \\ H \end{pmatrix}^{-1} \begin{pmatrix} x_i \\ x_d \end{pmatrix} = Lx_i + c_0, \qquad L = \begin{pmatrix} P \\ H \end{pmatrix}^{-1} \begin{pmatrix} I \\ 0 \end{pmatrix}, \quad c_0 = \begin{pmatrix} P \\ H \end{pmatrix}^{-1} \begin{pmatrix} 0 \\ c \end{pmatrix}$$

where c_0 represents the conserved quantities. We now write the dynamics for x_i as

$$\frac{dx_i}{dt} = P\frac{dx}{dt} = PNv(Lx_i + c_0, \theta) = N_r v_r(x_i, c_0, \theta), \tag{3.19}$$

where N_r is the *reduced stoichiometry matrix* and v_r is the rate vector with the conserved quantities separated out as constant parameters.

The reduced order dynamics in equation (3.19) represent the evolution of the independent species in the reaction. Given x_i, we can reconstruct the full set of species from the dynamics of the independent species by writing $x = Lx_i + c_0$. The vector c_0 represents the values of the conserved quantities, which must be specified in order to compute the values of the full set of species. In addition, since $x = Lx_i + c_0$, we have that

$$\frac{dx}{dt} = L\frac{dx_i}{dt} = LN_r v_r(x_i, c_0, p) = LN_r v(x, \theta),$$

which implies that

$$N = LN_r.$$

Thus, L also reconstruct the reduced stoichiometry matrix from the reduced space to the full space.

Example 3.8 (Enzyme kinetics)**.** Consider an enzymatic reaction

$$\mathrm{E} + \mathrm{S} \underset{d}{\overset{a}{\rightleftharpoons}} \mathrm{C} \xrightarrow{k} \mathrm{E} + \mathrm{P},$$

whose full dynamics can be written as

$$\frac{d}{dt}\begin{pmatrix} S \\ E \\ C \\ P \end{pmatrix} = \begin{pmatrix} -1 & 1 & 0 \\ -1 & 1 & 1 \\ 1 & -1 & -1 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} aE \cdot S \\ dC \\ kC \end{pmatrix}.$$

The conserved quantities are given by

$$H = \begin{pmatrix} 0 & 1 & 1 & 0 \\ 1 & -1 & 0 & 1 \end{pmatrix}.$$

The first of these is the total enzyme concentration $E_{\mathrm{tot}} = E + C$, while the second asserts that the concentration of product P is equal to the free enzyme concentration E minus the substrate concentration S. If we assume that we start with substrate concentration S_0, enzyme concentration E_{tot} and no product or bound enzyme, then the conserved quantities are given by

$$c = \begin{pmatrix} E + C \\ S - E + P \end{pmatrix} = \begin{pmatrix} E_{\mathrm{tot}} \\ S_0 - E_{\mathrm{tot}} \end{pmatrix}.$$

There are many possible choices for the set of independent species $x_i = Px$, but since we are interested in the substrate and the product, we choose P as

$$P = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

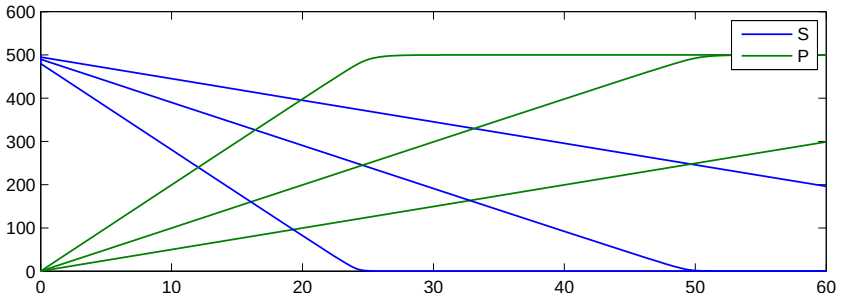


Figure 3.14: Enzyme dynamics. The simulations were carried out $a = d = 10$, $k = 1$, $S_0 = 500$ and $E_{\text{tot}} = 5, 1020$. The top plot shows the concentration of substrate S and product P, with the fastest case corresponding to $E_{\text{tot}} = 20$. The figures on the lower left zoom in on the substrate and product concentrations at the initial time and the figures on the lower right at one of the transition times.

Once P is chosen then we can compute

$$L = \begin{pmatrix} P \\ H \end{pmatrix}^{-1} \begin{pmatrix} I \\ 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \\ -1 & -1 \\ 0 & 1 \end{pmatrix}, \qquad c_0 = \begin{pmatrix} P \\ H \end{pmatrix}^{-1} \begin{pmatrix} 0 \\ c \end{pmatrix} = \begin{pmatrix} 0 \\ E_{\text{tot}} - S_0 \\ S_0 \\ 0 \end{pmatrix},$$

The resulting reduced order dynamics can be computed to be

$$\begin{aligned} \frac{d}{dt}\begin{pmatrix} S \\ P \end{pmatrix} &= \begin{pmatrix} -1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a(P + S + E_{\text{tot}} - S_0)S \\ d(-P - S + S_0) \\ k(-P - S + S_0) \end{pmatrix} \\ &= \begin{pmatrix} -a(P + S + E_{\text{tot}} - S_0)S - d(P + S - S_0) \\ k(S_0 - S - P) \end{pmatrix}. \end{aligned}$$

A simulation of the dynamics is shown in Figure 3.14. We see that the dynamics are very well approximated as being a constant rate of production until we exhaust the substrate (consistent with the Michaelis-Menten approximation).

∇

Metabolic control analysis

Metabolic control analysis (MCA) focuses on the study of the sensitivity of steady state concentrations and fluxes to changes in various system parameters. The basic concepts are equivalent to the sensitivity analysis tools described in Section 3.1, specialized to the case of reaction rate equations. In this section we provide a brief introduction to the key ideas, emphasizing the mapping between the general concepts and MCA terminology (as originally done by [43]).

Consider the reduced set of chemical reactions

$$\frac{dx_i}{dt} = N_r v_r(x_i, \theta) = N_r v(Lx_i + c_0, \theta).$$

We wish to compute the sensitivity of the equilibrium concentrations x_e and equilibrium fluxes v_e to the parameters θ. We start by linearizing the dynamics around an equilibrium point x_e. Defining $z = x - x_e$, $u = \theta - \theta_0$ and $f(z,u) = N_r v(x_e + z, \theta_0 + u)$, we can write the linearized dynamics as

$$\frac{dx}{dt} = Ax + Bu, \qquad A = \left(N_r \frac{\partial v}{\partial s} L\right), \quad B = \left(N_r \frac{\partial v}{\partial p}\right), \tag{3.20}$$

which has the form of a linear differential equation with state z and input u.

In metabolic control analysis, the following terms are defined:

$\bar{\epsilon}_\theta = \left.\dfrac{dv}{d\theta}\right\|_{x_e,\theta_o}$	$\bar{\epsilon}_\theta$ = flux control coefficients
$\bar{R}^x_\theta = \dfrac{\partial x_e}{\partial \theta} = \bar{C}^x \bar{\epsilon}_\theta$	$\bar{R}^x_\theta =$ $\bar{C}^x$ = concentration control coefficients
$\bar{R}^v_\theta = \dfrac{\partial v_e}{\partial \theta} = \bar{C}^v \bar{\epsilon}_\theta$	$\bar{R}^v_\theta =$ $\bar{C}^v$ = rate control coefficients

These relationships describe how the equilibrium concentration and equilibrium rates change as a function of the perturbations in the parameters. The two control matrices provide a mapping between the variation in the flux vector evaluated at equilibrium,

$$\left(\frac{\partial v}{\partial \theta}\right)_{x_e,\theta_0},$$

and the corresponding differential changes in the equilibrium point, $\partial x_e/\partial\theta$ and $\partial v_e/\partial\theta$. Note that

$$\frac{\partial v_e}{\partial \theta} \neq \left(\frac{\partial v}{\partial \theta}\right)_{x_e,\theta_0}.$$

The left side is the relative change in the equilibrium rates, while the right side is the change in the rate function $v(x,\theta)$ evaluated at an equilibrium point.

To derive the coefficient matrices $\bar{C}^x$ and $\bar{C}^v$, we simply take the linear equation (3.20) and choose outputs corresponding to s and v:

$$y_x = Ix, \qquad y_v = \frac{\partial v}{\partial x} Lx + \frac{\partial v}{\partial \theta} u.$$

Using these relationships, we can compute the transfer functions

$$H_x(s) = (sI - A)^{-1} B = \left[(sI - N_r \frac{\partial v}{\partial x} L)^{-1} N_r\right] \frac{\partial v}{\partial \theta},$$
$$H_v(s) = \frac{\partial v}{\partial s} L(sI - A)^{-1} B + \frac{\partial v}{\partial p} = \left[\frac{\partial v}{\partial x} L(sI - N_r \frac{\partial v}{\partial x} L)^{-1} N_r + I\right] \frac{\partial v}{\partial \theta}.$$

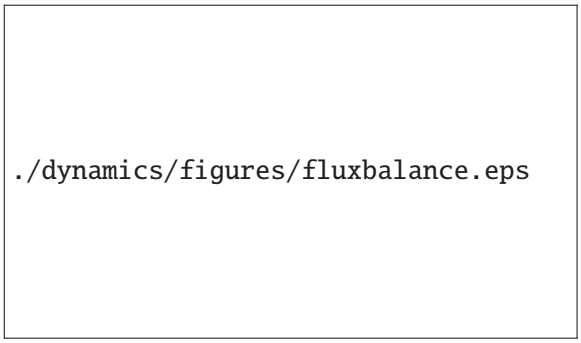


Figure 3.15: Flux balance analysis.

Classical metabolic control analysis considers only the equilibrium concentrations, and so these transfer functions would be evaluated at $x = 0$ to obtain the equilibrium equations.

These equations are often normalized by the equilibrium concentrations and parameter values, so that all quantities are expressed as fractional quantities. If we define

$$D^x = \text{diag}\{x_e\}, \qquad D^v = \text{diag}\{v(x_e, \theta_0)\}, \qquad D^\theta = \text{diag}\{\theta_0\},$$

then the normalized coefficient matrices (without the overbar) are given by

$$C^x = (D^x)^{-1}\bar{C}^x D^v, \qquad C^v = (D^v)^{-1}\bar{C}^v D^v,$$
$$R^x_\theta = (D^x)^{-1}\bar{R}^x_\theta D^\theta, \qquad R^v_\theta = (D^v)^{-1}\bar{R}^v_\theta D^\theta.$$

Flux balance analysis

Flux balance analysis is a technique for studying the relative rate of different reactions in a complex reaction system. We are most interested in the case where there may be multiple pathways in a system, so that the number of reactions m is greater than the number of species n. The dynamics

$$\frac{dx}{dt} = Nv(x, \theta)$$

thus have the property that the matrix N has more columns that rows and hence there are multiple reactions that can produce a given set of species. Flux balance is often applied to pathway analysis in metabolic systems to understand the limiting pathways for a given species and the the effects of changes in the network (e.g., through gene deletions) to the production capacity.

To perform a flux balance analysis, we begin by separating the reactions of the pathway into internal fluxes v_i versus exchanges flux v_e, as illustrated in Figure 3.15. The dynamics of the resulting system now be written as

$$\frac{dx}{dt} = Nv(x,\theta) = N\begin{pmatrix} v_i \\ v_e \end{pmatrix} = Nv_i(x,\theta) - b_e,$$

where $b_e = -Nv_e$ represents the effects of external fluxes on the species dynamics. Since the matrix N has more columns that rows, it has a *right* null space and hence there are many different internal fluxes that can produce a given change in species.

In particular, we are interested studying the steady state properties of the system. In this case, we have that $dx/dt = 0$ and we are left with an algebraic system

$$Nv_i = b_e.$$

Material to be completed. Review

3.4 Oscillatory Behavior

In addition to equilibrium behavior, a variety of cellular procesess involve oscillatory behavior in which the system state is constantly changing, but in a repeating pattern. Two examples of biological oscillations are the cell cycle and circadian rhythm. Both of these dynamic behaviors involve repeating changes in the concentrations of various proteins, complexes and other molecular species in the cell, though they are very different in their operation. In this section we discuss some of the underlying ideas for how to model this type of oscillatory behavior, focusing on those types of oscillations that are most common in biomolecular systems.

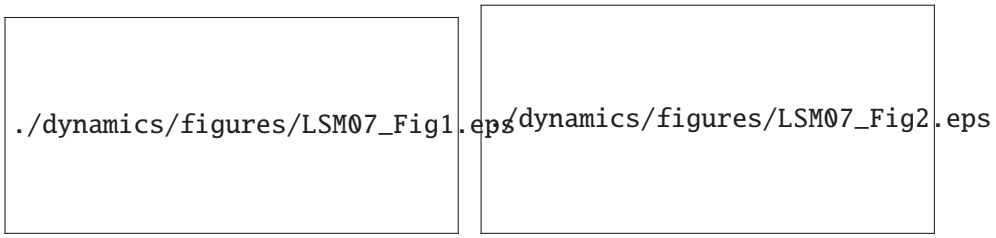
Biomolecular oscillators

Biological systems have a number of natural oscillatory processes that govern the behavior of subsystems and whole organisms. These range from internal oscillations within cells to the oscillatory nature of the beating heart to various tremors and other undesirable oscillations in the neuro-muscular system. At the biomolecular level, two of the most studied classes of oscillations are the cell cycle and circadian rhythm.

The cell cycle consists of a set "phases" that govern the duplication and division of cells into two new cells:

- G1 phase - gap phase, terminated by "G1 checkpoint"
- S phase - synthesis phase (DNA replication)
- G2 phase - gap phase, terminated by "G2 checkpoint"
- M - mitosis (cell division)

The cell goes through these stages in a cyclical fashion, with the different enzymes and pathways active in different phases. The cell cycle is regulated by many different proteins, often divided into two major classes. *Cyclins* are a class of proteins



(a) Overview of cell cycle (b) Molecular mechanisms

Figure 3.16: The *Caulobacter crescentus* cell cycle. (a) *Caulobacter* cells divide asymmetrically into a stalked cell, which is attached to a surface, and a swarmmer cell, that is motile. The swarmer cells can become stalked cells in a new location and begin the cell cycle anew. The transcriptional regulators CtrA, DnaA and GcrA are the primary factors that control the various phases of the cell cycle. (b) The genetic circuitry controlling the cell cycle consists of a large variety of regulatory mechanisms, described in more detail in the text. Figure obtained from [54] (permission TBD).

that sense environmental conditions internal and external to the cell and are also used to implement various logical operations that control transition out of the G1 and G2 phases. *Cyclin dependent kinases* (CDKs) are proteins that serve as "actuators" by turning on various pathways during different cell cycles.

An example of the control circuitry of the cell cycle for the bacterium *Caulobacter crescentus* (henceforth *Caulobacter*) is shown in Figure 3.16 [54]. This organism uses a variety of different biomolecular mechanisms, including transcriptional activation and represssion, positive autoregulation (CtrA), phosphotransfer and methlylation of DNA.

The cell cycle is an example of an oscillator that does not have a fixed period. Instead, the length of the individual phases and the transitioning of the different phases are determined by the environmental conditions. As one example, the cell division time for *E. coli* can vary between 20 minutes and 90 minutes due to changes in nutrient concentrations, temperature or other external factors.

A different type of oscillation is the highly regular pattern encoding in circadian rhythm, which repeats with a period of roughly 24 hours. The observation of circadian rhythms dates as far back as 400 BCE, when Androsthenes described observations of daily leaf movements of the tamarind tree [63]. There are three defining characteristics associated with circadian rhythm: (1) the time to complete one cycle is approximately 24 hours, (2) the rhythm is endogenously generated and self-sustaining and (3) the period remains relatively constant under changes in ambient temperature. Oscillations that have these properties appear in many different organisms, including micro-organisms, plants, insects and mammals. Some common features of the circuitry implementing circadian rhythms in these organisms is the combination of positive and negative feedback loops, often with the positive elements activating the expression of clock genes and the negative elements repress-

Figure 3.17: *Caption omitted pending permission.* (Figure and caption from [10])

ing the positive elements [10]. Figure 3.17 shows some of the different organisms in which circadian oscillations can be found and the primary genes responsible for different positive and negative factors.

Clocks, oscillators and limit cycles

To begin our study of oscillators, we consider a nonlinear model of the system described by the differential equation

$$\frac{dx}{dt} = f(x,u,\theta), \qquad y = h(x,\theta)$$

where $x \in \mathbb{R}^n$ represents the state of the system (typically concentrations of various proteins and other species and complexes), $u \in \mathbb{R}^q$ represents the external inputs, $y \in \mathbb{R}^p$ represents the (measured) outputs and $\theta \in \mathbb{R}^K$ represents the model parameters. We say that a solution $(x(t),u(t))$ is *oscillatory with period* T if $y(t+T) = y(t)$. For simplicity, we will often assume that $p = q = 1$, so that we have a single input and single output, but most of the results can be generalized to the multi-input, multi-output case.

There are multiple ways in which a solution can be oscillatory. One of the simplest is that the input $u(t)$ is oscillatory, in which case we say that we have a *forced oscillation*. In the case of a linear system, an input of the form $u(t) = A\sin\omega t$ then

we know already the output will be of the form $y(t) = M \cdot A\sin(\omega t + \phi)$ where M and ϕ represent the gain and phase of the system (at frequency ω). In the case of a nonlinear system, if the output is periodic then we can write it in terms of a set of harmonics,

$$y(t) = B_0 + B_1 \sin(\omega t + \phi_1) + B_2 \sin(2\omega t + \phi_2) + \cdots$$

The term B_0 represents the average value of the output (also called the bias), the terms B_i are the magnitudes of the ith harmonic and ϕ_i are the phases of the harmonics (relative to the input). The *oscillation frequency* ω is given by $\omega = 2\pi/T$ where T is the oscillation period.

A different situation occurs when we have no input (or a constant input) and still obtain an oscillatory output. In this case we say that the system has a *self-sustained oscillation*. This type of behavior is what is required for oscillations such as the cell cycle and circadian rhythm, where there is either no obvious forcing function or the forcing function is removed but the oscillation persists. If we assume that the input is constant, $u(t) = A_0$, then we are particularly interested in how the period T (or equivalently frequency ω), amplitudes B_i and phases ϕ_i depend on the input A_0 and system parameters θ.

To simplify our notation slightly, we consider a system of the form

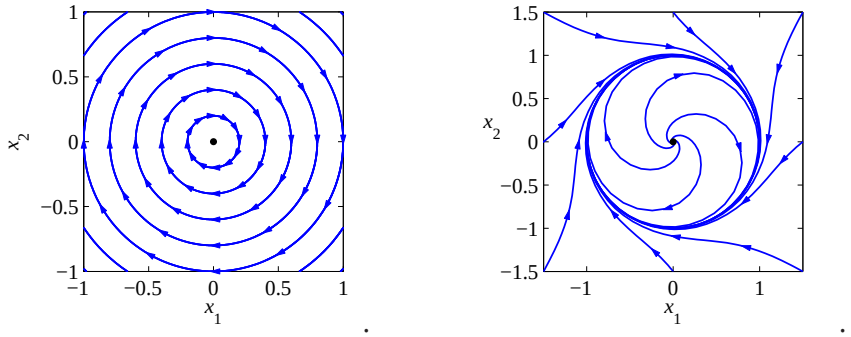
$$\frac{dx}{dt} = f(x, \theta), \qquad y = h(x, \theta) \tag{3.21}$$

where the input is ignored (or taken to be one of the constant parameters) in the analysis that follows. We have focused on the oscillatory nature of the output $y(t)$ thus far, but we note that if the states $x(t)$ are periodic then the output is as well, as this is the most common case. Hence we will often talk about the *system* being oscillatory, by which we mean that there is a solution for the dynamics in which the state satisfies $x(t + T) = x(t)$.

More formally, we say that a closed curve $\Gamma \in \mathbb{R}^n$ is an *orbit* if trajectories that start on Γ remain on Γ for all time and if Γ is not an equilibrium point of the system. As in the case of equilibrium points, we say that the orbit is *stable* if trajectories that start near Γ stay near Γ, *asymptotically stable* if in addition nearby trajectories approach Γ as $t \to \infty$ and *unstable* if it is not stable. The orbit Γ is periodic with period T if for any $x(t) \in \Gamma$, $x(t + T) = x(t)$.

There are many different types of periodic orbits that can occur in a system whose dynamics are modeled as in equation (3.21). A *harmonic oscillator* references to a system that oscillates around an equilibrium point, but does not (usually) get near the equilibrium point. The classical harmonic oscillator is a linear system of the form

$$\frac{d}{dt}\begin{pmatrix} 0 & \omega \\ -\omega & 0 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix},$$



(a) Linear harmonic oscillator (b) Nonlinear harmonic oscillator

Figure 3.18: Examples of harmonic oscillators.

whose solutions are given by

$$\begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix} = \begin{pmatrix} \cos\omega t & \sin\omega t \\ -\sin\omega t & \cos\omega t \end{pmatrix} \begin{pmatrix} x_1(0) \\ x_2(0) \end{pmatrix}.$$

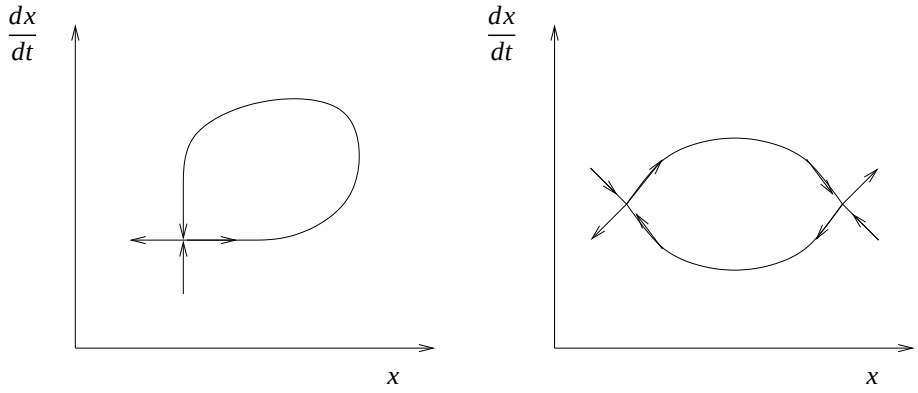
The frequency of this oscillation is fixed, but the amplitude depends on the values of the initial conditions, as shown in Figure 3.18. Note that this system has a single equilibrium point at $x = (0,0)$ and the eigenvalues of the equilibrium point have zero real part, so trajectories neither expand nor contract, but simply oscillate.

An example of a nonlinear harmonic oscillator is given by the equation

$$\frac{dx_1}{dt} = x_2 + x_1(1 - x_1^2 - x_2^2), \qquad \frac{dx_2}{dt} = -x_1 + x_2(1 - x_1^2 - x_2^2). \tag{3.22}$$

This system has an equilibrium point at $x = (0,0)$, but the linearization of this equilibrium point is unstable. The phase portrait in Figure 3.18b shows that the solutions in the phase plane converge to a circular trajectory. In the time domain this corresponds to an oscillatory solution. Mathematically the circle is called a *limit cycle*. Note that in this case, the solution for any initial condition approaches the limit cycle and the amplitude and frequency of oscillation "in steady state" (once we have reached the limit cycle) are independent of the initial condition.

A different type of oscillation can occur in nonlinear systems in which the equilibrium points are saddle points, having both stable and unstable eigenvalues. Of particular interest is the case where the stable and unstable orbits of one or more equilibrium points join together. Two such situations are shown in Figure 3.19. The figure on the left is an example of a *homoclinic orbit*. In this system, trajectories that start near the equilibrium point quickly diverge away (in the directions corresponding to the unstable eigenvalues) and then slowly return to the equilibrium point along the stable directions. If the initial conditions are chosen to be precisely on the homoclinic orbit Γ then the system slowly converges to the equilibrium



(a) Homoclinic orbit (b) Heteroclinic orbit

Figure 3.19: Homoclinic and heteroclinic orbits.

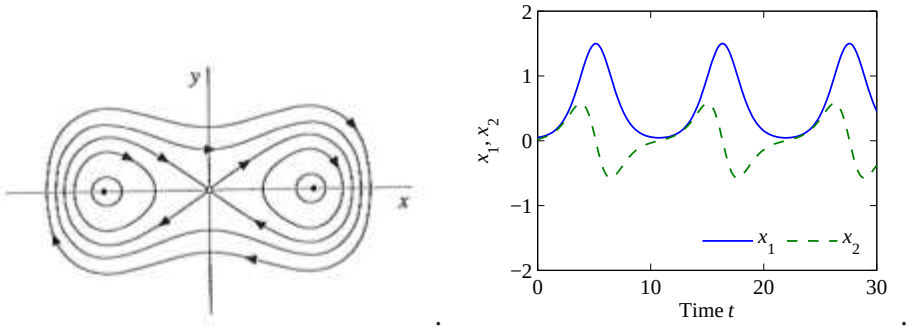
point, but in practice there are often disturbances present that will perturb the system off of the orbit and trigger a "burst" in which the system rapidly escapes from the equilibrium point and then slowly converges again.

A somewhat similar type of orbit is a *heteroclinic orbit*, in which the orbit connects two different equilibrium points, as shown in Figure 3.19b.

An example of a system with a homoclinic orbit is given by the system

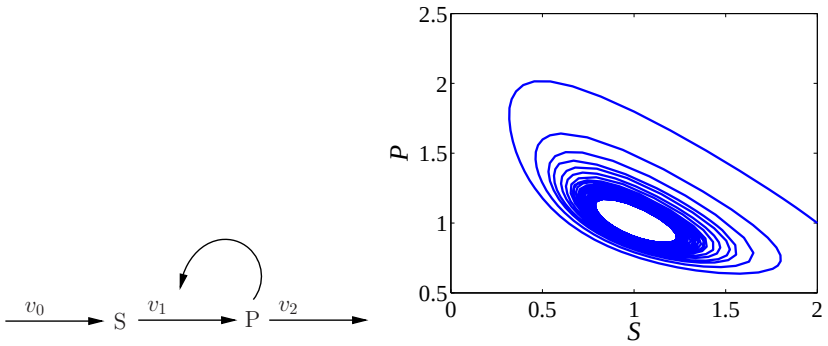
$$\frac{dx_1}{dt} = x_2, \qquad \frac{dx_2}{dt} = x_1 - x_1^3. \tag{3.23}$$

The phase portrait and time domain solutions are shown in Figure 3.20. In this system, there are periodic orbits both inside and outside the two homoclinic cycles (left and right). Note that the trajectory we have chosen to plot in the time domain has the property that it rapidly moves away from the equilibrium point and then slowly re-converges to the equilibrium point, before begin carried away



(a) (b)

Figure 3.20: Example of a homoclinic orbit.



(a) (b)

Figure 3.21: (a) The glycolysis pathway. "S" is a substrate, which is converted into product "P". This, in turn, is activating its own production by enhancing the rate v_2. (b) Oscillations in the glycolysis pathway. Parameters are $v_0 = 1$, $k_1 = 1$, and $k_2 = 1.00001$.

again. This type of oscillation, in which one slowly returns to an equilibrium point before rapidly diverging is often called a *relaxation oscillation*. Note that for this system, there are also oscillations that look more like the harmonic oscillator case described above, in which we oscillate around the unstable equilibrium points at $x = (\pm 1, 0)$.

Example 3.9 (Glycolytic oscillations)**.** Glycolysis is one of the principal metabolic networks involved in energy production. It is a sequence of enzyme-catalyzed reactions that coverts sugar into pyruvate, which is then further degraded to alcohol (in yeast fermentation) and lactic acid (in muscles) in anaerobic conditions, and ATP (the cell's major energy supply) is produced as a result. Both damped and sustained oscillations have been observed. Damped oscillations were first reported by [20] while sustained oscillations in yeast cell free extracts were observed when glucose-6-phosphate (G6P), fructose-6-phosphate (F6P) [39] or trehalose [75] were used as substrates.

Here, we introduce the fundamental motif that is known to be at the core of these oscillatory phenomenon. This is depicted in Figure 3.21 (a). A simple model for the system is given by the two differential equations

$$\frac{dS}{dt} = v_0 - v_1, \qquad \frac{dP}{dt} = v_1 - v_2,$$

in which

$$v_1 = S\,F(P), \quad F(P) = \frac{\alpha(P/K)^2}{1+(P/K)^2}, \qquad v_2 = k_2 P,$$

where $F(P)$ is the Hill function. Under the assumption that $K \gg P^2$, we have $F(P) \approx k_1 P^2$, in which we have defined $k_1 := \alpha/K$. This second order system admits a stable limit cycle under suitable parameter conditions (Figure 3.21b). ∇

The example above illustrates some of the types of questions we would like to answer for oscillatory systems. For example, under what parameter conditions do oscillations occur in the glycolytic system? How much can the parameter change before the limit cycle disappears? To analyze these sorts of questions, we need to introduce tools that allow to infer the existence and robustness of limit cycle behavior from a differential equation model. The objective of this section is to address these questions.

Consider the system $\dot{x} = f(x)$ and let $x(t, x_0)$ denote its solution starting at x_0 at time $t = 0$, that is, $\dot{x}(t, x_0) = f(x(t, x_0))$ and $x(0, x_0) = x_0$. We say that $x(t, x_0)$ is a *periodic solution* if there is $T > 0$ such that $x(t, x_0) = x(t + T, x_0)$ for all $t \in \mathbb{R}$. Here, we seek to answer two questions: (a) when does a system $\dot{x} = f(x)$ admit periodic solutions? (b) When are these periodic solutions stable or asymptotically stable?

In order to provide the main result to state the existence of a stable periodic solution, we need the concept of omega-limit set of a point p, denoted $\omega(p)$. Basically, the omega-limit set $\omega(p)$ denotes the set of all points to which the trajectory of the system starting from p tends as time approaches infinity. This is formally defined in the following definition.

Definition 3.1. A point $\bar{x} \in \mathbb{R}^n$ is called an *omega-limit point* of $p \in \mathbb{R}^n$ if there is a sequence of times $\{t_i\}$ with $t_i \to \infty$ for $i \to \infty$ such that $x(t_i, p) \to \bar{x}$ as $i \to \infty$. The *omega-limit set* of p, denoted $\omega(p)$, is the set of all omega-limit points of p.

The omega-limit set of a system has several relevant properties, among which the fact that it cannot be empty and that it must be a connected set.

Limit cycles in the plane

Before studying periodic behavior of systems in $\mathbb{R}^n$, we study the behavior of systems in $\mathbb{R}^2$ as several high dimensional systems can be often well approximated by systems in two dimensions by, for example, employing quasi-steady state approximations. For systems in $\mathbb{R}^2$, we will see that there are easy-to-check conditions that guarantee the existence of a limit cycle.

The first result that we next give provides a simple check to rule out periodic solutions for system in $\mathbb{R}^2$. Specifically, let $x \in \mathbb{R}^2$ and consider

$$\dot{x}_1 = f_1(x_1, x_2) \qquad \dot{x}_2 = f_2(x_1, x_2), \tag{3.24}$$

in which the functions $f_i : \mathbb{R}^2 \to \mathbb{R}^2$ for $i = 1, 2$ are smooth. Then, we have the following result:

Theorem 3.2 (Bendixson's criterion)**.** *If on a simply connected region $D \subset \mathbb{R}^2$ (i.e., there are no holes in it) the expression*

$$\frac{\partial f_1}{\partial x_1} + \frac{\partial f_2}{\partial x_2}$$

is not identically zero and does not change sign, then system (3.24) *has no closed orbits that lie entirely in D.*

Example 3.10. Consider the system

$$\dot{x}_1 = -x_2^3 + \delta x_1^3, \qquad \dot{x}_2 = x_1^3,$$

with $\delta \geq 0$. We can compute $\frac{\partial f_1}{\partial x_1} + \frac{\partial f_2}{\partial x_2} = 3\delta x_1^2$, which is positive in all $\mathbb{R}^2$ if $\delta \neq 0$. If $\delta \neq 0$, we can thus conclude from Bendixson's criterion that there are no periodic solutions. Investigate as an exercise what happens when $\delta = 0$. ∇

The following theorem, completely characterizes the omega-limit set of any point for a system in $\mathbb{R}^2$.

Theorem 3.3 (Poincarè-Bendixson)**.** *Let M be a bounded and closed positively invariant region for the system $\dot{x} = f(x)$ with $x \in$ (i.e., any trajectory that starts in M stays in M for all $t \geq 0$). Assume that there are finitely many steady states in M. Let $p \in M$, then one of the following possibilities holds for $\omega(p)$:*

(i) *$\omega(p)$ is a steady state;*

(ii) *$\omega(p)$ is a closed orbit;*

(iii) *$\omega(p)$ consists of a finite number of steady states and orbits, each starting (for $t = 0$) and ending (for $t \to \infty$) at one of the fixed points.*

This theorem has two important consequences:

1. If the system does not have steady states in M, since $\omega(p)$ is not empty, it must be a periodic solution;

2. If there is only one steady state in M and it is unstable and not a saddle (i.e., the eigenvalues of the linearization at the steady state are both positive), then $\omega(p)$ is a periodic solution.

We will employ this result in Chapter 6 to determine conditions under which the activator-repressor clock of Atkinson *et al.* [5] admits sustained oscillations.

Limit cycles in $\mathbb{R}^n$

The results above holds only for systems in two dimensions. However, there have been recent extensions of this theorem to systems with special structure in $\mathbb{R}^n$. In particular, we have the following result due to Hastings et al. (1977).

Theorem 3.4 (Hastings et al. 1977)**.** *Consider a system $\dot{x} = F(x)$, which is of the form*

$$\dot{x}_1 = f_1(x_n, x_1)$$
$$\dot{x}_j = f_j(x_{j-1}, x_j),\ 2 \leq j \leq n$$

on the set M defined by $x_i \geq 0$ for all i with the following inequalities holding in M:

(i) $\frac{\partial f_i}{\partial x_i} < 0$ *and* $\frac{\partial f_i}{\partial x_{i-1}} > 0$, *for* $2 \leq i \leq n$, *and* $\frac{\partial f_1}{\partial x_n} < 0$;

(ii) $f_i(0,0) \geq 0$ *and* $f_1(x_n, 0) > 0$ *for all* $x_n \geq 0$;

(iii) The system has a unique steady state $x^* = (x_1^*, ..., x_n^*)$ *in* M *such that* $f_1(x_n, x_1) < 0$ *if* $x_n > x_n^*$ *and* $x_1 > x_1^*$, *while* $f_1(x_n, x_1) > 0$ *if* $x_n < x_n^*$ *and* $x_1 < x_1^*$;

(iv) $\frac{\partial f_1}{\partial x_1}$ *is bounded above in* M.

Then, if the Jacobian of f at x^* has no repeated eigenvalues and has any eigenvalue with positive real part, then the system has a non-constant periodic solution in M.

This theorem states that for a system with cyclic structure in which the cycle "has negative gain", the instability of the steady state (under some technical assumption) is equivalent to the existence of a periodic solution. This theorem, however, does not provide information about whether the orbit is attractive or not, that is, of whether it is an omega-limit set of any point in M. This stability result is implied by a more recent theorem due to Mallet-Paret and Smith (1990), for which we provide a simplified statement as follows.

Theorem 3.5 (Mallet-Paret and Smith, 1990)**.** *Consider the system $\dot{x} = F(x)$ with the following cyclic feedback structure*

$$\begin{aligned} \dot{x}_1 &= f_1(x_n, x_1) \\ \dot{x}_j &= f_j(x_{j-1}, x_j), \quad 2 \leq j \leq n \end{aligned}$$

on a set M defined by $x_i \geq 0$ for all i with all trajectories starting in M bounded for $t \geq 0$. Then, the ω-limit set $\omega(p)$ of any point $p \in M$ can be one of the following:

(a) A steady state;

(b) A non-constant periodic orbit;

(c) A set of steady states connected by homoclinic or heteroclinic orbits.

As a consequence of the theorem, we have that for a system with cyclic feedback structure that admits one steady state only and at which the linearization has all eigenvalues with positive real part, the omega-limit set must be a periodic orbit.

Let for some $\delta_i \in \{1, -1\}$ be $\delta_i \frac{\partial f_i(x, x_{i-1})}{\partial x_{i-1}} > 0$ for all $0 \leq i \leq n$ and define $\Delta := \delta_1 \cdot ... \cdot \delta_n$. One can show that the sign of Δ is related to whether the system has one or multiple steady states.

In Chapter 6, we will apply these results in Chapter 6 to determine the parameter space that makes the repressilator [23] oscillate.

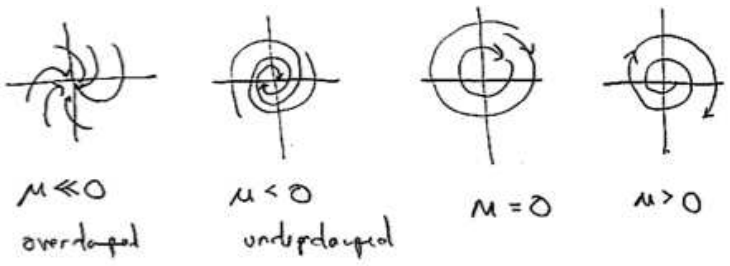


Figure 3.22: Phase portraits for a simple bifurcation.

3.5 Bifurcations

Another important property of nonlinear systems is how their behavior changes as the parameters governing the dynamics change. We can study this in the context of models by exploring how the location of equilibrium points, their stability, their regions of attraction and other dynamic phenomena, such as limit cycles, vary based on the values of the parameters in the model.

Parametric stability

Consider a differential equation of the form

$$\frac{dx}{dt} = F(x,\theta), \quad x \in \mathbb{R}^n, \theta \in \mathbb{R}^k, \tag{3.25}$$

where x is the state and θ is a set of parameters that describe the family of equations. The equilibrium solutions satisfy

$$F(x,\theta) = 0,$$

and as θ is varied, the corresponding solutions $x_e(\theta)$ can also vary. We say that the system (3.25) has a *bifurcation* at $\theta = \theta^*$ if the behavior of the system changes qualitatively at θ^*. This can occur either because of a change in stability type or a change in the number of solutions at a given value of θ.

As an example of a bifurcation, consider the linear system

$$\frac{dx_1}{dt} = x_2, \qquad \frac{dx_2}{dt} = -kx_1 - \mu x_2,$$

where $k > 0$ is fixed and θ is our bifurcation parameter. Figure 3.22 shows the phase portraits for different values of θ. We see that at $\theta = 0$ the system transitions from a single stable equilibrium point at the original to having an unstable equilibrium. Hence, as θ goes from negative to positive values, the behavior of the system changes in a significant way, indicating a bifurcation.

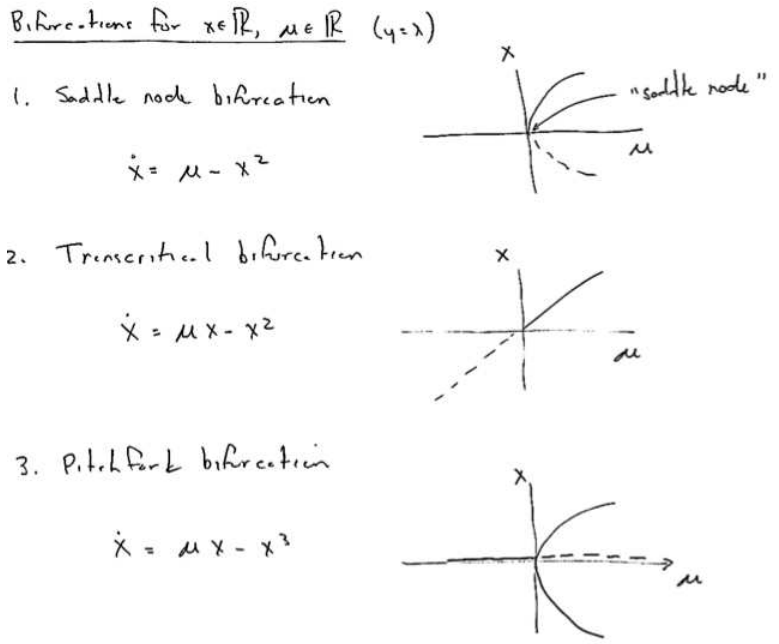
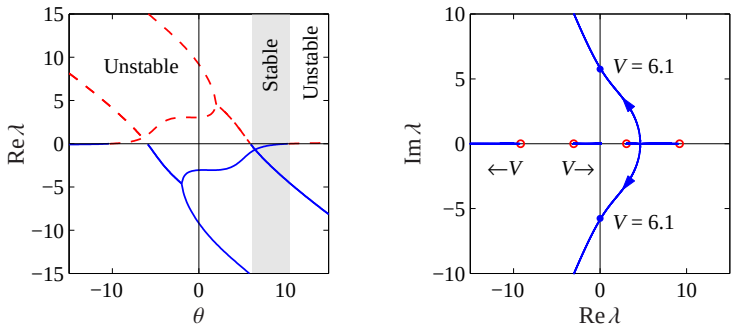


Figure 3.23: Bifurcation diagrams for some common bifurcations

A common way to visualize a bifurcation is through the use of a *bifurcation diagram*. To create a bifurcation diagram, we choose a function $y = h(x)$ such that the value of y at an equilibrium point has some useful meaning for the question we are studying. We then plot the value of $y_e = h(x_e(\theta))$ as a function of θ for all equilibria that exist for a given parameter value θ. By convention, we use dashed lines if the corresponding equilibrium point is unstable and solid lines otherwise. Figure 3.23 shows examples of some common bifurcation diagrams. Note that for some types of bifurcations, such as the pitchfork bifurcation, there exist values of θ where there is more than one equilibrium point. A system that exhibits this type of behavior is said to be *multistable*. A common case is that there are two stable equilibria, in which case the system is said to be *bistable*.

Another type of diagram that is useful in understanding parametric dependence is a *parametric stability diagram*, an example of which was shown in Figure **??**. In this type of diagram, we pick one or two (or sometimes three) parameters in the system and then analyze the stability type for the system over all possible combinations of those parameters. The resulting diagram shows those regions in parameter space where the system exhibits qualitatively different behaviors; an example is shown in Figure 3.24a.

A particular form of bifurcation that is very common when controlling linear systems is that the equilibrium remains fixed but the stability of the equilibrium



(a) Stability diagram (b) Root locus diagram

Figure 3.24: Stability plots a nonlinear system. The plot in (a) shows the real part of the system eigenvalues as a function of the parameter θ. The system is stable when all eigenvalues have negative real part (shaded region). The plot in (b) shows the locus of eigenvalues on the complex plane as the parameter θ is varied and gives a different view of the stability of the system. This type of plot is called a *root locus diagram*.

changes as the parameters are varied. In such a case it is revealing to plot the eigenvalues of the system as a function of the parameters. Such plots are called *root locus diagrams* because they give the locus of the eigenvalues when parameters change. An example is shown in Figure 3.24b. Bifurcations occur when parameter values are such that there are eigenvalues with zero real part. Computing environments such LabVIEW, MATLAB and Mathematica have tools for plotting root loci.

Parametric stability diagrams and bifurcation diagrams can provide valuable insights into the dynamics of a nonlinear system. It is usually necessary to carefully choose the parameters that one plots, including combining the natural parameters of the system to eliminate extra parameters when possible. Computer programs such as `AUTO`, `LOCBIF` and `XPPAUT` provide numerical algorithms for producing stability and bifurcation diagrams.

Hopf bifurcation

The bifurcations discussed above involved bifurcation of equilibrium points. Another type of bifurcation that can occur is that a system with an equilibrium point admits a limit cycle as a parameter is changed through a critical value. The Hopf bifurcation theorem provides a technique that is often used to understand whether a system admits a periodic orbit when some parameter is varied. Usually, such an orbit is a small amplitude periodic orbit that is present in the close vicinity of an unstable steady state.

Consider the system dependent on a parameter α:

$$\frac{dx}{dt} = g(x, \alpha), x \in \mathbb{R}^n, \ \alpha \in \mathbb{R},$$

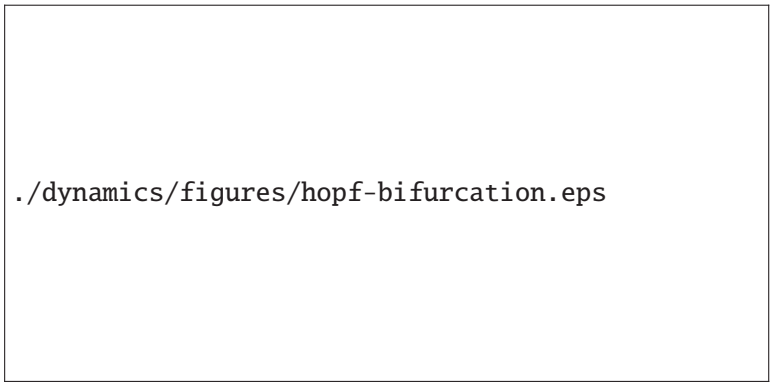


Figure 3.25: Hopf Bifurcation. On the left hand, as θ increases a stable limit cycle appears. On the right hand side, as θ increases a limit cycle appears but it is unstable.

and assume that at the steady state $\bar{x}$ corresponding to $\alpha = \bar{\alpha}$ (i.e., $g(\bar{x},\bar{\alpha}) = 0$), the linearization $\partial g/\partial x(\bar{x},\bar{\alpha})$ has a pair of (non zero) imaginary eigenvalues with the remaining eigenvalues having negative real parts. Define the new parameter $\theta := \alpha - \bar{\alpha}$ and re-define the system as

$$\frac{dx}{dt} = f(x,\theta) =: g(x,\theta+\bar{\alpha}),$$

so that the linearization $\partial f/\partial x(\bar{x},0)$ has a pair of (non zero) imaginary eigenvalues with the remaining eigenvalues having negative real parts. Denote by $\lambda(\theta) = \beta(\theta) + i\omega(\theta)$ the eigenvalue such that $\beta(0) = 0$. Then, if $\partial\beta/\partial\theta(0) \neq 0$ the system admits a small amplitude almost sinusoidal periodic orbit for θ small enough and the system is said to go through a Hopf bifurcation at $\theta = 0$. If the small amplitude periodic orbit is stable, the Hopf bifurcation is said *supercritical*, while if it is unstable it is said *subcritical*. Figure 3.25 shows diagrams corresponding to these bifurcations.

In order to determine whether a Hopf bifurcation is supercritical or subcritical, it is necessary to calculate a "curvature" coefficient, for which there are formulas (Marsden and McCrocken, 1976) and available bifurcation software, such as AUTO. In practice, it is often enough to calculate the value $\bar{\alpha}$ of the parameter at which Hopf bifurcation occurs and simulate the system for values of the parameter α close to $\bar{\alpha}$. If a small amplitude limit cycle appears, then the bifurcation must be supercritical.

Example 3.11 (Glycolytic oscillations)**.** Recalling the model (3.9) for the glycolytic oscillator, we ask whether such an oscillator goes through a Hopf bifurcation. In order to answer this question, we consider again the expression of the eigenvalues

$$\lambda_{1,2} = \frac{\mathrm{tr}(J) \pm \sqrt{\mathrm{tr}(J)^2 - 4\det(J)}}{2},$$

in which

$$\mathrm{tr}(J) = k_2 - k_1\left(\frac{v_0}{k_2}\right)^2 \quad \text{and} \quad \det(J) = k_1\left(\frac{v_0}{k_2}\right)^2.$$

The eigenvalues are imaginary if $\mathrm{tr}(J) = 0$, that is, if $k_1 = k_2^3/v_0^2$. Furthermore, the frequency of oscillations is given by $\omega = \sqrt{4\det(J)} = 4k_1(v_0/k_2)^2$. When $k_1 \approx k_2^3/v_0^2$, an approximately sinusoidal oscillation appears. When k_1 is large, the Hopf bifurcation theorem does not imply the existence of a periodic solution. This is because the Hopf theorem provides only local results. ∇

The Hopf bifurcation theorem is based on center manifold theory for nonlinear dynamical systems. For a rigorous treatment of Hopf bifurcation is thus necessary to study center manifold theory first, which is outside the scope of this text. For details, the reader is referred to standard text in dynamical systems [96, 37].

3.6 Model Reduction Techniques

The techniques that we have developed in this chapter can be applied to a wide variety of dynamical systems. However, many of the methods require significant computation and hence we would like to reduce the complexity of the models as much as possible before applying them. In this section, we review methods for doing such a reduction in the complexity of the models. Most of the techniques are based on the common idea that if we are interested in the slower time scale dynamics of a system, the fast time scale dynamics can be approximated by their equilibrium solutions. This idea was introduced in Chapter 2 in the context of reduced order mechanisms; we present a more mathematical analysis of such systems here.

Singular perturbation analysis

Singular perturbation techniques apply to systems that have processes that evolve on both fast and slow time scales and that can be written in a standard form, which we now introduce. Let $(x,y) \in D := D_x \times D_y \subset \mathbb{R}^n \times \mathbb{R}^m$ and consider the vector field

$$\begin{aligned} \frac{dx}{dt} &= f(x,y,\epsilon), & x(0) &= x_0 \\ \epsilon\frac{dy}{dt} &= g(x,y,\epsilon), & y(0) &= y_0 \end{aligned}$$

in which $0 < \epsilon \ll 1$ is a small parameter and both $f(x,y,0)$ and $g(x,y,0)$ are well defined. Since $\epsilon \ll 1$, the rate of change of y can be much larger than the rate of change of x, resulting in y dynamics that are much faster than the x dynamics. That is, this system has a slow time scale evolution (in x) and a fast time-scale evolution (in y), so that x is called the slow variable and y is called the fast variable.

If we are interested only in the slower time scale then the above system can be approximated (under suitable conditions) by the *reduced system*

$$\begin{aligned} \frac{d\bar{x}}{dt} &= f(\bar{x},\bar{y},0), \qquad \bar{x}(0) = x_0, \\ 0 &= g(\bar{x},\bar{y},0). \end{aligned}$$

Let $y = h(x)$ denote the locally unique solution of $g(x,y,0) = 0$. The manifold of (x,y) points where $y = h(x)$ is called the *slow manifold*. The *implicit function theorem* [60] shows that this solution exists whenever $\partial g/\partial y$ is non singular and that in such a case

$$\frac{dh}{dx} = -\frac{\partial g}{\partial y}^{-1}\frac{\partial g}{\partial x}.$$

We can now re-write the dynamics of x in the reduced system as

$$\frac{d\bar{x}}{dt} = f(\bar{x},h(\bar{x}),0), \qquad \bar{x}(0) = x_0.$$

We seek to determine under what conditions the solution $x(t)$ is "close" to the solution $\bar{x}(t)$ of the reduced system. This problem can be addressed by analyzing the fast dynamics, that is the dynamics of the system in the fast time scale $\tau = t/\epsilon$. In this case, we have that

$$\frac{dx}{d\tau} = \epsilon f(x,y,\epsilon), \qquad \frac{dy}{d\tau} = g(x,y,\epsilon), \qquad (x(0),y(0)) = (x_0,y_0),$$

so that when $\epsilon \ll 1$, $x(\tau)$ does not appreciably change. Therefore, the above system in the τ time scale can be well approximated by the system

$$\frac{dy}{d\tau} = g(x_0,y,0), \qquad y(0) = y_0,$$

in which x is "frozen" at the initial condition x_0. This system is usually referred to as the *boundary layer* system. For this system, the point $y = h(x_0)$ is an equilibrium point. Such an equilibrium point is asymptotically stable if $y(\tau)$ converges to $h(x_0)$ as $\tau \to \infty$. In this case, the solution $(x(t),y(t))$ of the original system approaches $(\bar{x}(t),h(\bar{x}(t)))$. This qualitative explanation is more precisely captured by the following theorem [51].

Theorem 3.6. *Assume that*

$$Real\left(\lambda\left(\frac{\partial}{\partial y}g(x,y)\Big|_{y=h(x)}\right)\right) < 0$$

uniformly for $x \in D_x$. Let the solution of the reduced system be uniquely defined for $t \in [0,t_f]$. Then, for all $t_b \in (0,t_f]$ there is a constant $\epsilon^* > 0$ and set $\Omega \subseteq D$ such that

$$\begin{aligned} x(t) - \bar{x}(t) &= O(\epsilon) \text{ uniformly for } t \in [0,t_f], \\ y(t) - h(\bar{x}(t)) &= O(\epsilon) \text{ uniformly for } t \in [t_b,t_f], \end{aligned}$$

provided $\epsilon < \epsilon^*$ and $(x_0,y_0) \in \Omega$.

Example 3.12 (Hill function)**.** In Section 2.1, we obtained the expression of the Hill function by making a quasi-steady state approximation on the dynamics of binding. Here, we illustrate how Hill function expressions can be derived by a formal application of singular perturbation. Specifically, consider the simple binding scenario of a transcription factor X with DNA promoter sites p. Assume that such a transcription factor is acting as an activator of the promoter and let Y be the protein expressed under promoter p. Assume further that X dimerizes before binding to promoter p. The reaction equations describing this system are given by

$$\mathrm{X}+\mathrm{X} \underset{k_2}{\overset{k_1}{\rightleftharpoons}} \mathrm{X}_2, \qquad \mathrm{X}_2+\mathrm{p} \underset{d}{\overset{a}{\rightleftharpoons}} \mathrm{C}, \qquad \mathrm{C} \xrightarrow{k_\mathrm{f}} \mathrm{m}_\mathrm{Y}+\mathrm{C},$$

$$\mathrm{m}_\mathrm{Y} \xrightarrow{\kappa} \mathrm{m}_\mathrm{Y}+\mathrm{Y}, \qquad \mathrm{m}_\mathrm{Y} \xrightarrow{\delta} \emptyset, \qquad \mathrm{Y} \xrightarrow{\gamma} \emptyset, \qquad p+C=p_{\mathrm{tot}}.$$

The corresponding differential equation model is given by

$$\begin{aligned}
\frac{dX_2}{dt} &= k_1X^2 - k_2X_2 - aX_2(p_{\mathrm{tot}}-C) + dC \\
\frac{dC}{dt} &= aX_2(p_{\mathrm{tot}}-C) - dC \\
\frac{dm_Y}{dt} &= k_\mathrm{f}C - \delta m_Y \\
\frac{dY}{dt} &= \kappa m_Y - \gamma Y.
\end{aligned}$$

Since all the binding reactions are much faster than mRNA and protein production and decay, we have that $k_2, d \gg k_\mathrm{f}, \kappa, \delta, \gamma$. Let $K_\mathrm{m} := k_2/k_1$, $K_\mathrm{d} := d/a$, $c := k_2/d$, and $\epsilon := \gamma/d$. Then, we can re-write the above system by using the substitutions

$$d = \frac{\gamma}{\epsilon},\ a = \frac{\gamma}{K_\mathrm{d}\epsilon},\ k_2 = c\frac{\gamma}{\epsilon},\ k_1 = c\frac{\gamma}{K_\mathrm{m}\epsilon},$$

so that we obtain

$$\begin{aligned}
\epsilon\frac{dX_2}{dt} &= c\frac{\gamma}{K_\mathrm{m}}X^2 - c\gamma X_2 - \frac{\gamma}{K_\mathrm{d}}X_2(p_{\mathrm{tot}}-C) + \gamma C \\
\epsilon\frac{dC}{dt} &= \frac{\gamma}{K_\mathrm{d}}X_2(p_{\mathrm{tot}}-C) - \gamma C \\
\frac{dm_Y}{dt} &= k_\mathrm{f}C - \delta m_Y \\
\frac{dY}{dt} &= \kappa m_Y - \gamma Y.
\end{aligned}$$

This system is in the standard singular perturbation form (3.6). As an exercise, the reader can verify that the slow manifold is locally asymptotically stable (see

Exercises). The slow manifold is obtained by setting $\epsilon = 0$ and determines X_2 and C as functions of X. These functions are given by

$$X_2 = \frac{X^2}{K_\mathrm{m}}, \qquad C = \frac{p_\mathrm{tot} X^2/(K_\mathrm{m}K_\mathrm{d})}{1 + X^2/(K_\mathrm{m}K_\mathrm{d})}.$$

As a consequence, the reduced system becomes

$$\begin{aligned} \frac{dm_Y}{dt} &= k_\mathrm{f} \frac{p_\mathrm{tot} X^2/(K_\mathrm{m}K_\mathrm{d})}{1 + X^2/(K_\mathrm{m}K_\mathrm{d})} - \delta m_Y \\ \frac{dY}{dt} &= \kappa m_Y - \gamma Y, \end{aligned}$$

which is the familiar expression for the dynamics of gene expression with an activator as derived in Section 2.1 and letting $\alpha = k_\mathrm{f} p_\mathrm{tot}$, we have that

$$F(X) = \alpha \frac{X^2/(K_\mathrm{m}K_\mathrm{d})}{1 + X^2/(K_\mathrm{m}K_\mathrm{d})}$$

is the standard Hill function expression.

. ∇

Example 3.13 (Enzymatic reaction)**.** Let's go back to the enzymatic reaction

$$\mathrm{E} + \mathrm{S} \underset{d}{\overset{a}{\rightleftharpoons}} \mathrm{C} \xrightarrow{k} \mathrm{E} + \mathrm{P},$$

in which E is an enzyme, S is the substrate to which the enzyme binds to form the complex C, and P is the product resulting from the modification of the substrate S due to the binding with the enzyme E. The corresponding system of differential equations is given by

$$\frac{dE}{dt} = -aE \cdot S + dC + kC, \qquad \frac{dC}{dt} = aE \cdot S - (d+k)C, \tag{3.26}$$

$$\frac{dS}{dt} = -aE \cdot S + dC, \qquad \frac{dP}{dt} = kC. \tag{3.27}$$

By considering that binding and unbinding reactions are much faster than the catalytic rates, mathematically expressed by $d \gg k$, we obtained before that approximately $dC/dt = 0$ and thus that $C = E_\mathrm{tot} S/(S + K_m)$, with $K_\mathrm{m} = (d+k)/a$ and $dP/dt = V_{max} S/(S + K_\mathrm{m})$ with $V_{max} = kE_\mathrm{tot}$. From this, it also follows that

$$\frac{dE}{dt} \approx 0 \text{ and } \frac{dS}{dt} \approx -\frac{dP}{dt}. \tag{3.28}$$

How good is this approximation? By applying the singular perturbation method, we will obtain a clear answer to this question. Specifically, define $K_\mathrm{d} := d/a$ and

take the system to standard singular perturbation form by defining the small parameter $\epsilon := k/d$, so that $d = k/\epsilon$, $a = k/(K_d\epsilon)$, and the system becomes

$$\epsilon\frac{dE}{dt} = -\frac{k}{K_d}E\cdot S + kC + \epsilon kC, \qquad \epsilon\frac{dC}{dt} = \frac{k}{K_d}E\cdot S - kC - \epsilon kC,$$
$$\epsilon\frac{dS}{dt} = -\frac{k}{K_d}E\cdot S + kC, \qquad \frac{dP}{dt} = kC.$$

One cannot directly apply singular perturbation theory on this system because one can verify from the linearization of the first three equations that the boundary layer dynamics are not locally exponentially stable since there are two zero eigenvalues. This is because the three variables E, S, C are not independent. Specifically, $E = E_{\text{tot}} - C$ and $S + C + P = S(0) = S_{\text{tot}}$, assuming that initially we have S in amount $S(0)$ and no amount of P and C in the system. Given these conservation laws, the system can be re-written as

$$\epsilon\frac{dC}{dt} = \frac{k}{K_d}(E_{\text{tot}} - C)\cdot(S_{\text{tot}} - C - P) - kC - \epsilon kC, \qquad \frac{dP}{dt} = kC.$$

Under the assumption made in the analysis of the enzymatic reaction that $S_{\text{tot}} \gg E_{\text{tot}}$, we have that $C \ll S_{\text{tot}}$ so that the equations finally become

$$\epsilon\frac{dC}{dt} = \frac{k}{K_d}(E_{\text{tot}} - C)\cdot(S_{\text{tot}} - P) - kC - \epsilon kC, \qquad \frac{dP}{dt} = kC.$$

One can verify (see Exercises) that in this system, the boundary layer dynamics is locally exponentially stable, so that setting $\epsilon = 0$ one obtains

$$\bar{C} = \frac{E_{\text{tot}}(S_{\text{tot}} - \bar{P})}{(S_{\text{tot}} - \bar{P}) + K_m} =: h(\bar{P})$$

and thus that the reduced system is given by

$$\frac{d\bar{P}}{dt} = V_{max}\frac{(S_{\text{tot}} - \bar{P})}{(S_{\text{tot}} - \bar{P}) + K_m}.$$

This system is the same as that obtained in Chapter 2. However, $dC(t)/dt$ and $dE(t)/dt$ are not close to zero as obtained earlier. In fact, from the conservation law $\bar{S} + \bar{C} + \bar{P} = S(0) = S_{\text{tot}}$, we obtain that $\frac{d\bar{S}}{dt} = -\frac{d\bar{P}}{dt} - \frac{d\bar{C}}{dt}$, in which now $\frac{d\bar{C}}{dt} = \frac{\partial h}{\partial P}(\bar{P})\cdot\frac{dP}{dt}$. Therefore

$$\frac{d\bar{S}}{dt} = -\frac{d\bar{P}}{dt}(1 + \frac{\partial h}{\partial P}(\bar{P})),\ \bar{S}(0) = S_{\text{tot}} - h(\bar{P}(0)) - \bar{P}(0) \tag{3.29}$$

and

$$\frac{d\bar{E}}{dt} = -\frac{d\bar{C}}{dt} = -\frac{\partial h}{\partial P}(\bar{P})\frac{d\bar{P}}{dt},\ E(0) = E_{\text{tot}} - h(\bar{P}(0)), \tag{3.30}$$

which are different from expressions (3.28).

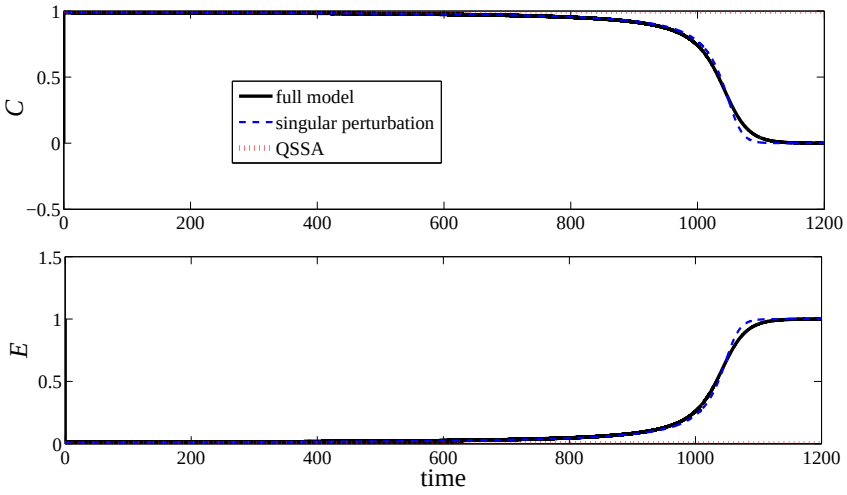


Figure 3.26: Simulation results for the enzymatic reaction comparing the approximations from singular perturbation and from the quasi-steady state approximation (QSSA). Here, we have $S_{\text{tot}} = 100$, $E_{\text{tot}} = 1$, $a = d = 10$, and $k = 0.1$. The full model is the one in equations (3.27).

These expressions are close to those in equation (3.28) only when $\partial h/\partial P(\bar{P})$ is small enough. In the plots of Figure 3.26, we show the time trajectories of the original system, of the Michaelis-Menten quasi-steady state approximation (QSSA), and of the singular perturbation approximation. In the full model (solid line in Figure 3.26), $E(t)$ starts from a unit concentration and immediately collapses to zero as the enzyme is all consumed to form the complex C by the substrate, which is in excess. Similarly, $C(t)$ starts from zero and immediately reaches the maximum possible value of one.

In the QSSA, both $E(t)$ and $C(t)$ are assumed to stabilize immediately to their (quasi) steady state and then stay constant. This is depicted by the dotted plots in Figure 3.26, in which $E(t)$ stays at zero for the whole time and $C(t)$ stays at one for the whole time. This approximation is fairly good as long as there is an excess of substrate. When the substrate concentration goes to zero as it is all converted to product, also the complex concentration C goes to zero (see solid line of Figure 3.26). At this time, the concentrations of complex and enzyme substantially change with time and the QSSA is unsatisfactory. By contrast, the reduced dynamics obtained from the singular perturbation approach well represent the dynamics of the full system even during this transient behavior. Hence, while the QSSA is a good approximation only as long as there is excess of substrate in the system, the reduced dynamics obtained by singular perturbation is a good approximation even when the substrate concentration goes to zero.

In Figure 3.27, we show the curve $C = h(P)$ (in red) and the trajectories of the

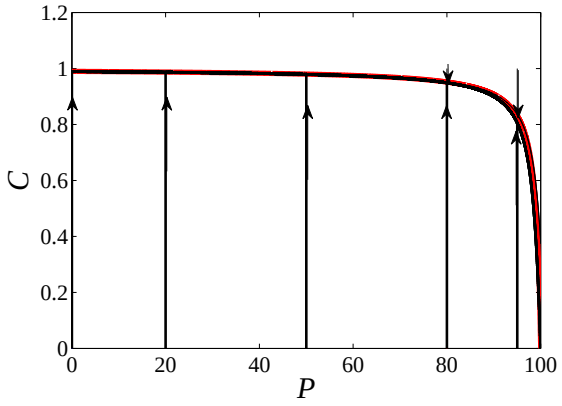


Figure 3.27: The slow manifold of the system $C = h(P)$ is shown in red. In black, we show the trajectories of the the full system. These trajectories collapse into an ϵ-neighbor of the slow manifold. Here, we have $S_{\text{tot}} = 100$, $E_{\text{tot}} = 1$, $a = d = 10$, and $k = 0.1$.

full system in black. All of the trajectories of the system immediately collapse into an ϵ-neighbor of the curve $C = h(P)$. From this plot, it is clear that $\partial h/\partial P$ is small as long as the product concentration P is small enough, which corresponds to a substrate concentration S large enough. This confirms that the QSSA is good only as long as there is excess of substrate S. ∇

Exercises

3.1 (BE 150, Winter 2011) **File missing:** ./dynamics/exercises/dual-activation

3.2 (BE 150, Winter 2011) **File missing:** ./dynamics/exercises/posfbk-cascade

3.3 (Frequency response of a phosphorylation cycle) Consider the model of a covalent modification cycle as illustrated in Chapter 2 in which the kinase Z is not constant, but it is produced and decays according to the reaction $\text{Z} \underset{\gamma}{\overset{u(t)}{\rightleftharpoons}}$. Let $u(t)$ be the input stimulus of the cycle and let X^* be the output. Determine the frequency response of X^* to u, determine its bandwidth, and make plots of it. What parameters can be used to tune the bandwidth?

3.4 (Design for robustness) Consider a one-step reaction model for a phosphorylation cycle as seen in Homework 1, in which the input stimulus is the time-varying concentration of kinase $Z(t)$. When found in the cellular environment, this cycle is subject to possible interactions with other cellular components, such as the non-specific or specific binding of X* to target sites, to noise due to stochasticity of the cellular environment, and to other cross-talk phenomena. We will come back to

these "disturbances" later during the course. For now, we can think of these disturbances as acting like an aggregate rate of change on the output protein X*, which we call $d(t)$. Hence, we can model the "perturbed" cycle by

$$\dot{X}^* = Z(t)k_1 X_{\text{tot}}\left(1 - \frac{X^*}{X_{\text{tot}}}\right) - k_2 Y_{\text{tot}} X^* + d(t),$$

which is the same as you found in Homework 1, except for the presence of the disturbance $d(t)$. Assume that you can tune all the parameters in this system (we will see later that this is actually possible to large extent by suitably fabricating genetic circuits). Can you tune these parameters so that the response of $X^*(t)$ to $d(t)$ is arbitrarily attenuated while the response of $X^*(t)$ to $Z(t)$ remains arbitrarily large? If yes, explain how these parameters should be tuned to reach this design objective and justify your answer through a careful mathematical reasoning using the tools introduced in class.

3.5 (Adaptation) Show that the equation of the sniffer (3.15) can be taken into the standard integral feedback form through a suitable change of coordinates.

3.6 (Design limitations) This problem is meant to have you think about possible trade-offs and limitations that are involved in any realistic design question (we will come back to this when we start design). Here, we examine this through the open loop and negative feedback transcriptional component seen in class (see Figure 3-8 in the Lecture Notes). Specifically, we want to compare the robustness of these two topologies to cellular noise, crosstalk, and other cellular interactions. As performed in Problem 1, we model these phenomena as a time-varying disturbance affecting the production rate of mRNA m and protein P. To slightly simplify the problem, we focus only on disturbances affecting the production of protein. The open loop model becomes

$$\dot{m}_{\text{P}} = \alpha_0 - \delta m_{\text{P}} \qquad \dot{P} = \kappa m_{\text{P}} - \gamma P + d(t)$$

and the negative feedback system becomes

$$\dot{m}_{\text{P}} = \alpha_0 + \frac{\alpha}{1 + (P/K)^n} - \delta m_{\text{P}} \qquad \dot{P} = \beta m_{\text{P}} - \gamma P + d(t).$$

Answer the following questions:

(a) After performing linearization about the equilibrium point, determine analytically the frequency response of P to d for both systems.

(b) Sketch the magnitude plot of this response by hand for both systems, compare them, and determine what happens as κ and α increase (note: if your calculations are correct, you should find that what really matters for the negative feedback system is the product $\alpha\kappa$, which we can view as the *feedback*

gain). So, is increasing the feedback gain to arbitrarily large values the best strategy to decrease the sensitivity of the system to the disturbance? Comment.

(c) Pick parameter values and use Matlab to draw Bode plots as the feedback gain increases and validate your predictions of (b). (Suggested parameters: $\delta = 1$, $\gamma = 1$, $K = 1$, $n = 1$, $\alpha\kappa = \{1, 10, 100, 1000, ...\}$). Note: in Matlab, once you have determined the matrices A, B, C, and D for the linearization, you can just do: `SYS=ss(A,B,C,D); bode(SYS)` and the Bode plot will pop up.

(d) Investigate the answer to (c) when you have $\delta = 20$, that is, the timescale of the mRNA dynamics becomes faster than that of the protein dynamics. What does change with respect to what you found in (c)? Note: when δ increases you are reducing the (phase) lag within the negative feedback loop...

(e) When δ is at least 10 times larger than γ, you can approximate the m dynamics to the quasi-steady state. So, the two above systems can be reduced to one differential equation each for the protein concentration P. For these two reduced systems, determine analytically the frequency response to d and use it to find out whether arbitrarily increasing the feedback gain is a good strategy to decrease the sensitivity of response to the disturbance.

3.7 (Bendixson criterion) Consider the possible circuit topologies of Figure 3.28, in which A and B are transcriptional components. Model each transcriptional component by a first order system, in which you have approximated the mRNA dynamics at the quasi-steady state. Hence, each topology will be represented by a dynamical system in the plane $\mathbb{R}^2$. Use Bendixson criterion to rule out topologies that cannot give rise to closed orbits.

3.8 (Two gene oscillator) Consider the feedback system composed of two genes expressing proteins A (activator) and R (repressor), in which we denote by A, R, m_A, and m_R, the concentrations of the activator protein, the repressor protein, the mRNA for the activator protein, and the mRNA for the repressor protein, respectively. The ODE model corresponding to this system is given by

$$\frac{dm_A}{dt} = \frac{\alpha}{1+(R/K_1)^n} - \delta m_A \qquad \frac{dm_R}{dt} = \frac{\alpha(A/K_2)^m}{1+(A/K_2)^m} - \delta m_R$$
$$\frac{dA}{dt} = \kappa m_A - \gamma A \qquad \frac{dR}{dt} = \kappa m_R - \gamma R.$$

Determine parameter conditions under which this system admits a stable limit cycle. Validate your finding through simulation.

3.9 (Goodwin oscillator) Consider the simple set of reactions

$$\mathrm{X}_1 \xrightarrow{k} X_2 \xrightarrow{k} X_3 .... \xrightarrow{k} \mathrm{X}_\mathrm{n}.$$

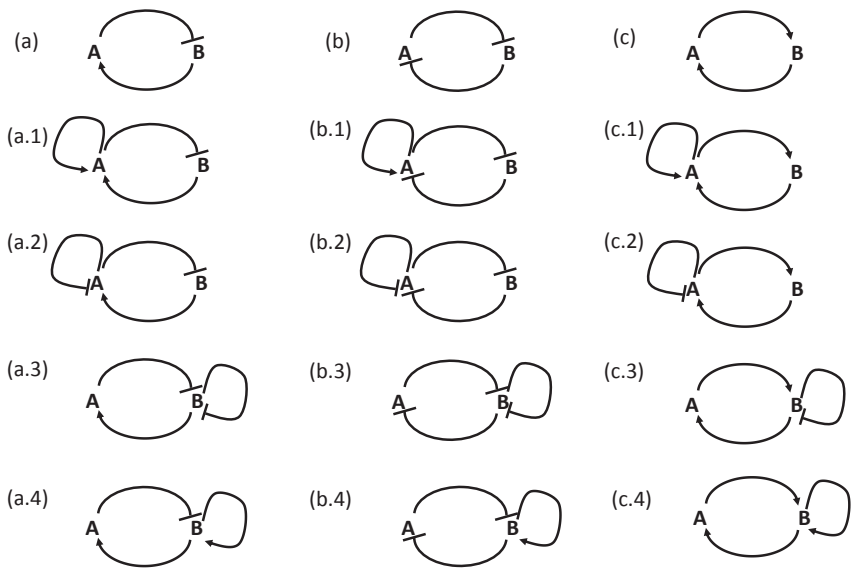


Figure 3.28: Circuit topologies with two components (proteins): A and B.

Assume further that X_n is a transcription factor that represses the production of protein X_1 through transcriptional regulation (assume simple binding of X_1 to DNA). Neglecting the mRNA dynamics of X_1, write down the ODE model of this system and determine conditions on the length n of the cascade for which the system admits a stable limit cycle. Validate your finding through simulation.

3.10 (Phosphorylation via singular perturbation) Consider again the model of a covalent modification cycle as illustrated in Chapter 2 in which the kinase Z is not constant, but it is produced and decays according to the reaction $\mathrm{Z} \underset{u(t)}{\overset{\gamma}{\rightleftharpoons}} \emptyset$.

(a) Consider that $a, d \gg k, \gamma, u(t)$ and employ singular perturbation with small parameter, for example, $\epsilon = \gamma/d$ to obtain the approximated dynamics of $Z(t)$ and $X^*(t)$. How is this different from the result obtained in Exercise 2.8? Explain.

(b) Simulate these approximated dynamics when $u(t)$ is a periodic signal with frequency ω and compare the responses of Z of this approximated dynamics to those obtained in Exercise 2.8 as you change ω. What do you observe? Explain.

3.11 (Hill function via singular perturbation) Show that the slow manifold of the

following system is asymptotically stable:

$$\epsilon\frac{dX_2}{dt} = c\frac{\gamma}{K_{\mathrm{m}}}X^2 - c\gamma X_2 - \frac{\gamma}{K_{\mathrm{d}}}X_2(p_{\mathrm{tot}} - C) + \gamma C, \qquad \frac{dm_Y}{dt} = \alpha C - \delta m_Y,$$
$$\epsilon\frac{dC}{dt} = \frac{\gamma}{K_{\mathrm{d}}}X_2(p_{\mathrm{tot}} - C) - \gamma C, \qquad \frac{dY}{dt} = \beta m_Y - \gamma Y.$$

3.12 (Enzyme dynamics via singular perturbation) Show that the slow manifold of the following system is asymptotically stable:

$$\epsilon\frac{dC}{dt} = \frac{k}{K_{\mathrm{d}}}(E_{\mathrm{tot}} - C)\cdot(S_{\mathrm{tot}} - P) - kC - \epsilon kC, \qquad \frac{dP}{dt} = kC.$$

3.13 (BE 150, Winter 2011; Based on Alon 4.6—Shaping the pulse) Consider a situation where X in an I1-FFL begins to be produced at time t=0, so that the level of protein X gradually increases. The input signal S_x and S_y are present throughout.

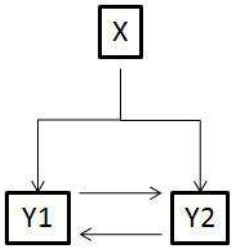
(a) How does the pulse shape generated by the I1-FFL depend on the thresholds K_{xz}, K_{xy}, and K_{yz}, and on β, the production rate of protein X? (i.e. How does increasing or decreasing these parameters change the height or position of the pulse peak, the slope of the rise of the pulse, etc?)

(b) Analyze a set of genes $Z_1, Z_2, ..., Z_n$, all regulated by the same X and Y in I1-FFLs. Design thresholds such that the genes are turned ON in the rising phase of the pulse in a certain temporal order and turned OFF in the declining phase of the pulse with the same order.

(c) Design thresholds such that the turn-OFF order is opposite the turn-ON order. Plot the resulting dynamics.

3.14 (BE 150, Winter 2011; Based on Alon 5.6—Bi-fan dynamics) Consider a bi-fan in which activators X_1 and X_2 regulate genes Z_1 and Z_2. The input signal of X_1, S_X2, appears at time t=0 and vanishes at time t=D. The input signal of X_2, S_X2, appears at time t=D/2 and vanishes at t=2D. Plot the dynamics of the promoter activity of Z_1 and Z_2 given that the input functions of Z_1 and Z_2 are AND and OR logic, respectively.

3.15 (BE 150, Winter 2011; Based on Alon 6.1—Memory in the regulated-feedback network motif) Transcription factor X activates transcription factor Y_1 and Y_2. Y_1 and Y_2 mutually activate each other. The input function at the Y_1 and Y_2 promoters is an OR gate (Y_2 is activated when either X or Y_1 binds the promoter). At time t=0, X begins to be produced from an initial concentration of X=0. Initially $Y_1 = Y_2 = 0$. All production rates are $\beta = 1$ and degradation rates are $\alpha = 1$. All of the activation thresholds are K=0.5. At time t=3, production of X stops.



(a) Plot the dynamics of X, Y_1, Y_2. What happens to Y_1 and Y_2 after X decays away?

(b) Consider the same problem, but now Y_1 and Y_2 repress each other and X activates Y_1 and represses Y_2. At time t=0, X begins to be produced and the initial levels are $X = 0, Y_1 = 0, Y_2 = 1$. At time t=3, X production stops. Plot the dynamics of the system. What happens after X decays away?

3.16 (BE 150, Winter 2011; Repressilator) Simulate the following simplified version of the repressilator:

$$\frac{dm_1}{dt} = \frac{k_p}{1+(\frac{p_3}{K_M})^n} - k_{mdeg}m_1 \qquad \frac{dp_1}{dt} = k_{trans}m_1 - k_{pdeg}p_1$$

$$\frac{dm_2}{dt} = \frac{k_p}{1+(\frac{p_1}{K_M})^n} - k_{mdeg}m_2 \qquad \frac{dp_2}{dt} = k_{trans}m_2 - k_{pdeg}p_2$$

$$\frac{dm_3}{dt} = \frac{k_p}{1+(\frac{p_2}{K_M})^n} - k_{mdeg}m_3 \qquad \frac{dp_3}{dt} = k_{trans}m_3 - k_{pdeg}p_3$$

(a) Simulate the system using the following parameters: $k_p = 0.5, n = 2, K_M = 40, k_{mdeg} = 0.0058, k_{pdeg} = 0.0012, k_{trans} = 0.116$.

(b) Suppose the protein half-life suddenly decreases by half. Which parameter(s) will change and how? Simulate what happens. What if the protein half-life is doubled? How do these two changes affect the oscillatory behavior?

(c) Now assume that there is leakiness in the transcription process. How does the system's ODE change? Simulate the system with a small leakiness (say, 5e-3) and comment on how it affects the oscillatory behavior.

3.17 (BE 150, Winter 2011; Glycolytic oscillations) In almost all living cells, glucose is broken down into the cell's energy currency, ATP, via the glycolysis pathway. Glycolysis is autocatalytic in the sense that ATP must first be consumed in the early steps before being produced later and oscillations in glycolytic metabolites have been observed experimentally. We will look at a minimal model of glycolysis:

$$\frac{dX}{dt} = \frac{2Vy^a}{1+y^h} - kx \qquad \frac{dY}{dt} = (q+1)kx - q\frac{2Vy^a}{1+y^h} - 1$$

Note that this system has been normalized such that $Y_{ss} = 1$.

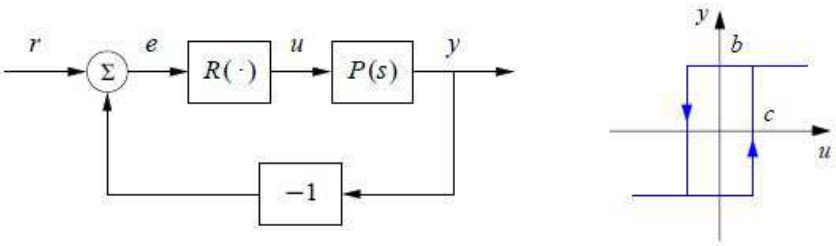
(a) While a system may have the potential to oscillate, the behavior still depends on the parameter values. The glycolysis system undergoes multiple *bifurcations* as the parameters are varied. Using linear stability analysis, find the parameter conditions where the system is stable vs. unstable. Next, find the conditions where the system has eigenvalues with nonzero imaginary parts.

(b) Let $q=k=V=1$. Find the relationship between h and a where the system is stable or not. Draw the stability diagram and mark the regions where the system is stable vs. unstable. In the same plot, mark the regions where the system has eigenvalues with nonzero imaginary parts.

(c) Let $q=k=V=1$. Choose h and a such that the eigenvalues are unstable and have nonzero imaginary parts. Use these parameter values and simulate the nonlinear system in MATLAB. Sketch the time response of the system starting with initial condition X(0) =1.2, Y(0) = 0.5 (you may use MATLAB or sketch by hand). Comment on what you see compared to what linear stability analysis told you about the system.

3.18 (BE 150, Winter 2011) Finding limit cycles for nonlinear systems and understanding how changes in parameters affect the amplitude and period of the oscillation is difficult to do in analytical form. A graphical technique that gives some insight into this problem is the use of *describing functions*, which is described in *Feedback Systems*, Section 9.5. In this problem we will use describing functions for a simple feedback system to approximate the amplitude and frequency of a limit cycle in analytical form.

Consider the system with the block diagram shown below. The block R is a relay



with hysteresis whose input/output response is shown on the right and the process transfer function is $P(s) = e^{-s\tau}/s$. Use describing function analysis to determine frequency and amplitude of possible limit cycles. Simulate the system and compare with the results of the describing function analysis.

3.19 (BE 150, Winter 2011) In this problem we will compare the model with single methylation site vs. double methylation sites. The model with a single methylation

site is given by:

$$\frac{d(X+X*)}{dt} = V_R R - \frac{V_B B X*}{K+X*}$$

where the *activity* is given by $A = X*$. The model with two methylation sites is given by

$$\frac{d(X_2+X_2*)}{dt} = \frac{RV_R X_1}{X_1+X_0} - BV_B X_2*$$

$$\frac{d(X_1+X_1*)}{dt} = BV_B X_2 * + \frac{RV_R X_0}{X_1+X_0} - \frac{RV_R X_1}{X_1+X_0} - BV_B X_1*$$

$$\frac{dX_0}{dt} = -\frac{RV_R X_0}{X_0+X_1} + BV_B X_1*$$

and the activity is given by $A = X_1 * + X_2*$. Let $K = 10, V_R R = 1, V_B B = 2$. Derive the parameter sensitivities of the activities ($\frac{dA}{dp_i}$) for both the single and double methylation models. Comment on which parameter each model is most robust and most sensitive to.

3.20 (BE 150, Winter 2011) Consider a toy model of protein production:

$$\frac{dm}{dt} = f(p) - \delta m \qquad\qquad \frac{dp}{dt} = g(p) - \gamma p$$

(a) Assume that there is transcriptional self-regulation ($f(p) = \frac{\alpha}{K+p^n}$). We now know that the mRNA transcription process and thus we want to understand the sensitivity with respect to the mRNA transcription rate α_0. Compute the transfer function from α to p. Plot this transfer function for $\alpha = 0.002, \beta_0 = 0.1, \delta = 0.005, \gamma = 0.001, K = 0.002$. Compare it with the transfer function from α_0 to p without regulation ($f(p) = \alpha_0 = 0.001$). (Note: As a reminder on how to compute these transfer functions, see BFS chapter 3 page 3-11).

(b) Now assume that there is no transcriptional regulation ($f(p) = \alpha_0$) but there is translational self-regulation such that $g(p) = \frac{\beta m}{K+p^n}$. Computer the transfer function from α_0 to p when $\beta = 0.2$. Compare again with the case with no regulation.

3.21 (BE 150, Winter 2011) Consider a simple model of chemotaxis:

$$\frac{dX_m}{dt} = k_R R + k^f(L) X_m^* - k^r X_m$$

$$\frac{dX_m^*}{dt} = -k_B B^p \frac{X_m^*}{K_{X_m^*} + X*_m} - k^f(L) X_m^* + k^r X_m$$

where X_m is the concentration of methylated receptor complex, and X_m^* is the concentration of activated, methylated receptor complex. Ligand concentration enters into the equation through the rate $k^f(L)$. In this model, *CheR* (R) and *CheB*^{P} (B^P) concentrations are constant. (BFS, Section 5)

(a) Pick parameter values such that $k_B B^p > k_R R$ and plot the dynamics, doubling the ligand concentration at time t=20. Compare to figure 5.12 in BFS.

(b) Now assume that CheR no longer acts in saturation. Rederive the dynamics and plot. Comment on how this assumption affects adaptation.

Chapter 4
Stochastic Modeling and Analysis

In this chapter we explore stochastic behavior in biomolecular systems, building on our preliminary discussion of stochastic modeling in Section 2.1. We begin by reviewing the various methods for modeling stochastic processes, including the chemical master equation (CME), the chemical Langevin equation (CLE) and the Fokker-Planck equation (FPE). Given a stochastic description, we can then analyze the behavior of the system using a variety of stochastic simulation and analysis tools.

Prerequisites. This chapter makes use of a variety of topics in stochastic processes that are not covered in AM08. Readers should have a good working knowledge of basic probability and some exposure to simple stochastic processes (e.g., Brownian motion), at the level of the material presented in Appendix **??** (drawn from [68]).

4.1 Stochastic Modeling of Biochemical Systems

Biomolecular systems are inherently noisy due to the random nature of molecular reactions. When the concentrations of molecules are high, the deterministic models we have used in the previous chapters provide a good description of the dynamics of the system. However, if the molecular counts are low then it is often necessary to explictly account for the random nature of events. In this case, th chemical reactions in the cell can be modeled as a collection of stochastic events corresponding to chemical reactions between species, including binding and unbinding of molecules (such as RNA polymerase and DNA), conversion of one set of species into another, and enzymatically controlled covalent modifications such as phosphorylation. In this section we will briefly survey some of the different representations that can be used for stochastic models of biochemical systems, following the material in the textbooks by Phillips *et al.* [72], Gillespie [29] and Van Kampen [49].

Statistical mechanics

At the core of many of the reactions and multi-molecular interactions that take place inside of cells is the chemical physics associated with binding between two molecules. One way to capture some of the properties of these interactions is through the use of statistical mechanics and thermodynamics.

As described briefly already in Chapter 2, the underlying representation for both statistical mechanics and chemical kinetics is to identify the appropriate microstates of the system. A microstate corresponds to a given configuration of the components (species) in the system relative to each other and we must enumerate all possible configurations between the molecules that are being modeled.

In statistical mechanics, we model the configuration of the cell by the probability that system is in a given microstate. This probability can be calculated based on the energy levels of the different microstates. Consider a setting in which our system is contained within a reservoir. Let E_r represent the energy in the resevoir, E_s the energy in the system and $E_{\text{tot}} = E_r + Es$ the total (conserved) energy. Given two different energy levels $E_s^{(1)}$ and $E_s^{(2)}$ for the system of interest, let $W_r(E_{\text{tot}} - E_s^{(i)})$ be the number of possible microstates of the reservoir with energy $E_r = E_{\text{tot}} - E_s^{(i)}$, $i = 1, 2$. The laws of statistical mechanics state that the ratio of probabilities of being at the energy levels $E_s^{(1)}$ and $E_s^{(2)}$ is given by the ratio of number of possible states of the reservoir:

$$\frac{\mathbb{P}(E_s^{(1)})}{\mathbb{P}(E_s^{(2)})} = \frac{W_r(E_{\text{tot}} - E_s^{(1)})}{W_r(E_{\text{tot}} - E_s^{(2)})}. \tag{4.1}$$

Defining the entropy of the system as $S = k_B \ln W$, where k_B is Boltmann's constant, we can rewrite equation (4.1) as

$$\frac{W_r(E_{\text{tot}} - E_s^{(1)})}{W_r(E_{\text{tot}} - E_s^{(2)})} = \frac{e^{S_r(E_{\text{tot}} - E_s^{(1)})/k_B}}{e^{S_r(E_{\text{tot}} - E_s^{(2)})/k_B}}.$$

We now approximate $S_r(E_{\text{tot}} - E_s)$ in a Taylor series expansion around E_{tot}, under the assumption that $E_r \gg E_s$:

$$S_r(E_{\text{tot}} - E_s) \approx S_r(E_{\text{tot}}) - \frac{\partial S_r}{\partial E} E_s.$$

From the properties of thermodynamics, if we hold the volume and number of molecules constant, then we can define the temperature as

$$\left.\frac{\partial S}{\partial E}\right|_{V,N} = \frac{1}{T}$$

and we obtain

$$\frac{\mathbb{P}(E_s^{(1)})}{\mathbb{P}(E_s^{(2)})} = \frac{e^{-E_s^{(1)}/k_B T}}{e^{-E_s^{(2)}/k_B T}}.$$

This implies that

$$\mathbb{P}(E_s^{(q)}) \propto e^{-E_s^{(q)}/(k_B T)}$$

and hence the probability of being in a microstate q is given by

$$\mathbb{P}(q) = \frac{1}{Z} e^{-E_q/(k_B T)}, \tag{4.2}$$

where we have written E_q for the energy of the microstate and Z is a normalizing factor, known as the *partition function*, defined by

$$Z = \sum_{q \in Q} e^{-E_q/(k_B T)}.$$

By keeping track of those microstates that correspond to a given system state (also called a macrostate), we can compute the overall probability that a given macrostate is reached.

In order to determine the energy levels associated with different microstates, we will often make use of the *free energy* of the system. Consider an elementary reaction $\mathrm{A} + \mathrm{B} \rightleftharpoons \mathrm{AB}$. Let E be the energy of the system, taken to be operating at pressure P in a volume V. The *enthalpy* of the system is defined as $H = E + PV$ and the *Gibbs free energy* is defined as $G = H - TS$ where T is the temperature of the system and S is its entropy (defined above). The change in bond energy due to the reaction is given by

$$\Delta H = \Delta G + T\Delta S,$$

where the Δ represents the change in the respective quantity. $-\Delta H$ represents the amount of heat that is absorbed from the reservoir, which then affects the entropy of the reservoir.

Derivation to be added later. Review

The resulting formula for the probability of being in a microstate q is given by

$$\mathbb{P}(q) = \frac{1}{Z} e^{-\Delta G/k_B T}.$$

Example 4.1 (Transcription factor binding)**.** Suppose that we have a transcription factor R that binds to a specific target region on a DNA strand (such as the promoter region upstream of a gene). We wish to find the probability P_{bound} that the transcription factor will be bound to this location as a function of the number of transcription factor molecules n_{R} in the system. If the transcription factor is a repressor, for example, knowing $P_{\text{bound}}(n_{\mathrm{R}})$ will allow us to calculate the likelihood of transcription occurring.

To compute the probability of binding, we assume that the transcription factor can bind non-specifically to other sections of the DNA (or other locations in the cell) and we let N_{ns} represent the number of such sites. We let E_{bound} represent the free energy associated with R bound to its specified target region and E_{ns} represent the free energy for R in any other non-specific location, where we assume that $E_{\text{bound}} < E_{\text{ns}}$. The microstates of the system consist of all possible assignments of the n_{R} transcription factors to either a non-specific location or the target region of the DNA. Since there is only one target site, there can be at most one transcription factor attached there and hence we must count all of the ways in which either zero or one molecule of R are attached to the target site.

If none of the n_R copies of R are bound to the target region then these must be distributed between the N_{ns} non-specific locations. Each bound protein has energy E_{ns}, so the total energy for any such configuration is $n_R E_{ns}$. The number of such combinations is $\binom{N_{ns}}{n_R}$ and so the contribution to the partition function from these microstates is

$$Z_{ns} = \binom{N_{ns}}{n_R} e^{-n_R E_{ns}/(k_B T)} = \frac{N_{ns}!}{n_R!(N_{ns}-n_R)!} e^{-n_R E_{ns}/(k_B T)}$$

For the microstates in which one molecule of R is bound at a target site and the other $n_R - 1$ molecules are at the non-specific locations, we have a total energy of $E_{bound} + (n_R - 1)E_{ns}$ and $\binom{N_{ns}}{(n_R-1)}$ possible such states. The resulting contribution to the partition function is

$$Z_{bound} = \frac{N_{ns}!}{(n_R-1)!(N_{ns}-n_R+1)!} e^{-(E_{bound}-(n_R-1)E_{ns})/(k_B T)}.$$

The probability that the target site is occupied is now computed by looking at the ratio of the Z_{bound} to $Z = Z_{ns} + Z_{bound}$. After some basic algebraic manipulations, it can be shown that

$$P_{bound}(n_R) = \frac{\left(\frac{n_R}{N_{ns}-n_R+1}\right)\exp[-(E_{bound}+E_{ns})/(k_B T)]}{1+\left(\frac{n_R}{N_{ns}-n_R+1}\right)\exp[-(E_{bound}+E_{ns})/(k_B T)]}.$$

If we assume that $N_{ns} \gg n_R$ then $N_{ns} - n_R + 1 \approx N_{ns}$, and we can write

$$P_{bound}(n_R) \approx \frac{k n_R}{1 + k n_R}, \quad \text{where} \quad k = \frac{1}{N_{ns}} \exp[-(E_{bound} - E_{ns})/(k_B T)].$$

As we would expect, this says that for very small numbers of repressors, P_{bound} is close to zero, while for large numbers of repressors, $P_{bound} \to 1$. The point at which we get a binding probability of 0.5 is when $n_R = 1/k$, which depends on the relative binding energies and the number of non-specific binding sites. ∇

Example 4.2 (Combinatorial promoter)**.** A combinatorial promoter is a region of DNA in which multiple transcription factors can bind and influence the subsequent binding of RNA polymerase. Combinatorial promoters appear in a number of natural and engineered circuits and represent a mechanism for creating switch-like behavior, for example by having a gene that controls expression of its own transcription factors.

One method to model a combinatorial promoter is to use the binding energies of the different combinations of proteins to the operator region, and then compute the probability of being in a given promoter state given the concentration of each of the transcription factors. Table 4.1 shows the possible states of a notional promoter that has two operator regions—one that binds a repressor protein R and another

Table 4.1: Configurations for a combinatorial promoter with an activator and a repressor. Each row corresponds to a specific macrostate of the promoter in which the listed molecules are bound to the target region. The relative energy of state compared with the ground state provides a measure of the likelihood of that state occurring, with more negative numbers corresponding to more energetically favorable configurations.

State	OR1	OR2	Prom	E_q (ΔG)	Comment
S_1	–	–	–	0	No binding (ground state)
S_2	–	–	RNAP	−5	RNA polymerase bound
S_3	R	–	–	−10	Repressor bound
S_4	–	A	–	−12	Activator bound
S_5	–	A	RNAP	−15	Activator and RNA polymerase

that binds an activator protein A. As indicated in the table, the promoter has three (possibly overlapping) regions of DNA: OR1 and OR2 are binding sites for the repressor and activator proteins, and Prom is the location where RNA polymerase binds. (The individual labels are primarily for bookkeeping purposes and may not correspond to physically separate regions of DNA.)

To determine the probabilities of being in a given macrostate, we must compute the individual microstates that occur at a given concentrations of repressor, activator and RNA polymerase. Each microstate corresponds to an individual set of molecules binding in a specific configuration. So if we have n_R repressor molecules, then there is one microstate corresponding to *each* different repressor molecule that is bound, resulting in n_R individual microstates. In the case of configuration S_5, where two different molecules are bound, the number of combinations is given by the product of the numbers of individual molecules, $n_\mathrm{A} \cdot n_\mathrm{RNAP}$, reflecting the possible combinations of molecules that can occupy the promoter sites. The overall partition function is given by summing up the contributions from each microstate:

$$Z = e^{-E_0/(k_B T)} + n_\mathrm{RNAP}\, e^{-E_\mathrm{RNAP}/(k_B T)} + n_\mathrm{R}\, e^{-E_\mathrm{R}/(k_B T)} + n_\mathrm{A}\, e^{-E_\mathrm{A}/(k_B T)} + n_\mathrm{A} n_\mathrm{RNAP}\, e^{-E_\mathrm{A:RNAP}/(k_B T)}. \quad (4.3)$$

The probability of a given macrostate is determined using equation (2.2). For example, if we define the promoter to be "active" if RNA polymerase is bound to the DNA, then the probability of being in this macrostate as a function of the various molecular counts is given by

$$\begin{aligned} P_\mathrm{active}(n_\mathrm{R}, n_\mathrm{A}, n_\mathrm{RNAP}) &= \frac{1}{Z}\left(n_\mathrm{RNAP}\, e^{-E_\mathrm{RNAP}/(k_B T)} + n_\mathrm{A}\, n_\mathrm{RNAP} e^{-E_\mathrm{A:RNAP}/(k_B T)}\right) \\ &= \frac{k_\mathrm{A:RNAP}\, n_\mathrm{A} + k_\mathrm{RNAP}}{1 + k_\mathrm{RNAP} + k_\mathrm{R}\, n_\mathrm{R} + (k_\mathrm{A} + k_\mathrm{A:RNAP}) n_\mathrm{A}}, \end{aligned}$$

where

$$k_X = e^{-(E_X - E_0)/(k_B T)}.$$

From this expression we see that if $n_R \gg n_A$ then P_{active} tends to 0 while if $n_A \gg n_R$ then P_{active} tends to 1, as expected. ∇

Chemical master equation (CME)

The statistical physics model we have just considered gives a description of the *steady state* properties of the system. In many cases, it is clear that the system reaches this steady state quickly and hence we can reason about the behavior of the system just by modeling the free energy of the system. In other situations, however, we care about the transient behavior of a system or the dynamics of a system that does not have an equilibrium configuration. In these instances, we must extend our formulation to keep track of how quickly the system transitions from one microstate to another, known as the *chemical kinetics* of the system.

To model these dynamics, we return to our enumeration of all possible microstates of the system. Let $P(q,t)$ represent the probability that the system is in microstate q at a given time t. Here q can be any of the very large number of possible microstates for the system, which for chemical reaction systems we can represent in terms of a vector consisting of the number of molecules of each species that is present. We wish to write an explicit expression for how $P(q,t)$ varies as a function of time, from which we can study the stochastic dynamics of the system.

We begin by assuming we have a set of M reactions $\mathrm{R_j}$, $j = 1,\ldots,M$, with ξ_j representing the change in state associated with reaction $\mathrm{R_j}$. Specifically, ξ_j is given by the jth column of the stoichiometry matrix N. The *propensity function* defines the probability that a given reaction occurs in a sufficiently small time step dt:

$$a_j(q,t)dt \quad = \quad \text{Probability that reaction } \mathrm{R_j} \text{ will occur between time } t \text{ and time } t+dt \text{ given that the microstate is } q.$$

The linear dependence on dt relies on the fact that dt is chosen sufficiently small. We will typically assume that a_j does not depend on the time t and write $a_j(q)dt$ for the probability that reaction j occurs in state q.

Using the propensity function, we can compute the distribution of states at time $t+dt$ given the distribution at time t:

$$\begin{aligned} P(q,t+dt) &= P(q,t)\Big(1 - \sum_{j=1}^{M} a_j(q)dt\Big) + \sum_{j=1}^{M} P(q-\xi_j)a_j(q-\xi_j)dt \\ &= P(q,t) + \sum_{j=1}^{M}\Big(a_j(q-\xi_j)P(q-\xi_j,t) - a_j(q)P(q,t)\Big)dt. \end{aligned} \tag{4.4}$$

Since dt is small, we can take the limit as $dt \to 0$ and we obtain the *chemical master equation* (CME):

$$\frac{\partial P}{\partial t}(q,t) = \sum_{j=1}^{M} \Big(a_j(q-\xi_j) P(q-\xi_j,t) - a_j(q) P(q,t) \Big) \tag{4.5}$$

This equation is also referred to as the *forward Kolmogorov equation* for a discrete state, continuous time random process.

Despite its complexity, the master equation does capture many of the important details of the chemical physics of the system and we shall use it as our basic representation of the underlying dynamics. As we shall see, starting from this equation we can then derive a variety of alternative approximations that allow us to answer specific equations of interest.

The key element of the master equation is the propensity function $a_j(q,t)$, which governs the rate of transition between microstates. Although the detailed value of the propensity function can be quite complex, its functional form is often relatively simple. In particular, for a unimolecular reaction of the form $\mathrm{A} \to B$, the propensity function is proportional to the number of molecules of A that are present:

$$a_j(q,t) = k_j n_A. \tag{4.6}$$

This follows from the fact that each reaction is independent and hence the likelihood of a reaction happening depends directly on the number of copies of A that are present.

Similarly, for a bimolecular reaction, we have that the likelihood of a reaction occurring is proportional to the product of the number of molecules of each type that are present (since this is the number of independent reactions that can occur) and inversely proportional to the volume Ω. Hence, for a reaction of the form $\mathrm{A} + \mathrm{B} \longrightarrow \mathrm{C}$ we have

$$a_j(q,t) = \frac{k_j}{\Omega} n_A n_B. \tag{4.7}$$

The rigorous verification of this functional form is beyond the scope of this text, but roughly we keep track of the likelihood of a single reaction occurring between A and B and then multiply by the total number of combinations of the two molecules that can react ($n_\mathrm{A} \cdot n_\mathrm{B}$).

A special case of a bimolecular reaction occurs when $\mathrm{A} = \mathrm{B}$, so that our reaction is given by $\mathrm{A} + \mathrm{A} \to B$. In this case we must take into account that a molecule cannot react with itself, and so the propensity function is of the form

$$a_i(q,t) = \frac{1}{2} \frac{k_i}{\Omega} n_A (n_A - 1). \tag{4.8}$$

The term $n_A(n_A - 1)$ represents the number of ways that two molecules can be chosen from a collection of n_A identical molecules.

Table 4.2: Examples of propensity functions for some common cases [31]. Here we take r_a and r_b to be the effective radii of the molecules, $m^* = m_a m_b/(m_a + m_b)$ is the reduced mass of the two molecules, Ω is the volume over which the reaction occurs, T is temperature, k_B is Boltzmann's constant and n_A, n_B are the numbers of molecules of A and B present.

Reaction type	Propensity function coefficient, k_i
Reaction occurs if molecules "touch"	$\left(\frac{8k_BT}{\pi m^*}\right)^{1/2}\pi(r_a+r_b)^2$
Reaction occurs if molecules collide with energy ϵ	$\left(\frac{8k_BT}{\pi m^*}\right)^{1/2}\pi(r_a+r_b)^2\cdot e^{-\epsilon/k_BT}$
Steady state transcription factor	$P_{\text{bound}}k_{\text{oc}}n_{\text{RNAP}}$

Note that the use of the parameter k_i in the propensity functions above is intentional since it corresponds to the reaction rate parameter that is present in the reaction rate equation model. The factor of Ω for biomolecular reactions models the fact that the propensity of a biomolecular reaction occurring depends explicitly on the volume in which the reaction takes place.

Although it is tempting to extend the formula for a biomolecular reaction to the case of more than two species being involved in a reaction, usually such reactions actually involve combinations of bimolecular reactions, e.g.:

$$\mathrm{A}+\mathrm{B}+\mathrm{C}\longrightarrow \mathrm{D} \qquad \Longrightarrow \qquad \mathrm{A}+\mathrm{B}\longrightarrow \mathrm{AB} \quad \mathrm{AB}+\mathrm{C}\longrightarrow \mathrm{D}$$

This more detailed description reflects that fact that it is extremely unlikely that three molecules will all come together at precisely the same instant, versus the much more likely possibility that two molecules will initially react, followed be a second reaction involving the third molecule.

The propensity functions for these cases and some others are given in Table 4.2.

Example 4.3 (Repression of gene expression)**.** We consider a simple model of repression in which we have a promoter that contains binding sites for RNA polymerase and a repressor protein R. RNA polymerase only binds when the repressor is absent, after which it can undergo an isomerization reaction to form an open complex and initiate transcription. Once the RNA polymerase begins to create mRNA, we assume the promoter region is uncovered, allowing another repressor or RNA polymerase to bind.

The following reactions describe this process:

$$\begin{aligned}
\mathrm{R}_1&: \quad \mathrm{R}+\mathrm{DNA}\longrightarrow \mathrm{R{:}DNA}\\
\mathrm{R}_2&: \quad \mathrm{R{:}DNA}\longrightarrow \mathrm{R}+\mathrm{DNA}\\
\mathrm{R}_3&: \quad \mathrm{RNAP}+\mathrm{DNA}\longrightarrow \mathrm{RNAP{:}DNA^{C}}\\
\mathrm{R}_4&: \quad \mathrm{RNAP{:}DNA^{C}}\longrightarrow \mathrm{RNAP}+\mathrm{DNA}\\
\mathrm{R}_5&: \quad \mathrm{RNAP{:}DNA^{C}}\longrightarrow \mathrm{RNAP{:}DNA^{O}}\\
\mathrm{R}_6&: \quad \mathrm{RNAP{:}DNA^{O}}\longrightarrow \mathrm{RNAP}+\mathrm{DNA}+\mathrm{mRNA},
\end{aligned}$$

where RNAP:DNA^{c} represents the closed complex and RNAP:DNA^{o} represents the open complex. The states for the system depend on the number of molecules of each species that are present. If we assume that we start with n_R repressors and n_{RNAP} RNA polymerases, then the possible states for our system are given by

State	DNA	R	RNAP	R:DNA	RNAP:DNA^{c}	RNAP:DNA^{o}
q_1	1	n_R	n_{RNAP}	0	0	0
q_2	0	$n_R - 1$	n_{RNAP}	1	0	0
q_3	0	n_R	$n_{RNAP} - 1$	0	1	0
q_4	0	n_R	$n_{RNAP} - 1$	0	0	1

Note that we do not keep track of each individual repressor or RNA polymerase molecule that binds to the DNA, but simply keep track of whether they are bound or not.

We can now rewrite the chemical reactions as a set of transitions between the possible microstates of the system. Assuming that all reactions take place in a volume Ω, we use the propensity functions for unimolecular and bimolecular reactions to obtain:

$$\begin{aligned}
&\mathrm{R}_1: \quad \mathrm{q}_1 \longrightarrow \mathrm{q}_2; \quad a_1(q_1) = (k_1/\Omega)\, n_R && \mathrm{R}_2: \quad \mathrm{q}_2 \longrightarrow \mathrm{q}_1; \quad a_2(q_2) = k_2 \\
&\mathrm{R}_3: \quad \mathrm{q}_1 \longrightarrow \mathrm{q}_3; \quad a_3(q_1) = (k_3/\Omega)\, n_{RNAP} && \mathrm{R}_4: \quad \mathrm{q}_3 \longrightarrow \mathrm{q}_1; \quad a_4(q_3) = k_4 \\
&\mathrm{R}_5: \quad \mathrm{q}_3 \longrightarrow \mathrm{q}_4; \quad a_5(q_3) = k_5 && \mathrm{R}_6: \quad \mathrm{q}_4 \longrightarrow \mathrm{q}_1; \quad a_6(q_4) = k_6
\end{aligned}$$

The chemical master equation can now be written down using the propensity functions for each reaction:

$$\frac{d}{dt}\begin{pmatrix} P(q_1,t) \\ P(q_2,t) \\ P(q_3,t) \\ P(q_4,t) \end{pmatrix} = \begin{pmatrix} -(k_1/\Omega)n_R - (k_3/\Omega)n_{RNAP} & k_2 & k_4 & k_6 \\ (k_1/\Omega)n_R & -k_2 & 0 & 0 \\ (k_3/\Omega)n_{RNAP} & 0 & -k_4 - k_5 & 0 \\ 0 & 0 & k_5 & -k_6 \end{pmatrix} \begin{pmatrix} P(q_1,t) \\ P(q_2,t) \\ P(q_3,t) \\ P(q_4,t) \end{pmatrix}.$$

The initial condition for the system can be taken as $P(q,0) = (1,0,0,0)$, corresponding to the state q_1. A simulation showing the evolution of the probabilities is shown in Figure 4.1.

The equilibrium solution for the probabilities can be solved by setting $\dot{P} = 0$, which yields:

$$\begin{aligned}
P_e(q_1) &= \frac{k_2 k_4 \Omega (k_4 + k_5)}{k_1 k_6 n_R (k_4 + k_5) + k_2 k_3 n_{RNAP}(k_5 + k_6) + k_2 k_6 \Omega (k_4 + k_5)} \\
P_e(q_2) &= \frac{k_1 k_6 n_R (k_4 + k_5)}{k_1 k_6 n_R (k_4 + k_5) + k_2 k_3 n_{RNAP}(k_5 + k_6) + k_2 k_6 \Omega (k_4 + k_5)} \\
P_e(q_3) &= \frac{k_2 k_3 k_6 n_{RNAP}}{k_1 k_6 n_R (k_4 + k_5) + k_2 k_3 n_{RNAP}(k_5 + k_6) + k_2 k_6 \Omega (k_4 + k_5)} \\
P_e(q_4) &= \frac{k_2 k_3 k_5 n_{RNAP}}{k_1 k_6 n_R (k_4 + k_5) + k_2 k_3 n_{RNAP}(k_5 + k_6) + k_2 k_6 \Omega (k_4 + k_5)}
\end{aligned}$$

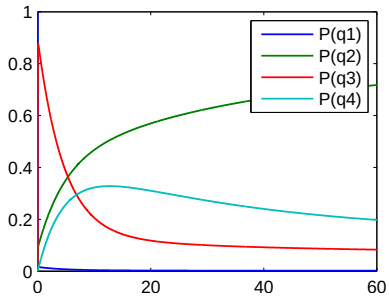


Figure 4.1: Numerical solution of chemical master equation for simple repression model.

We see that the functional dependencies are similar to the case of the combinatorial promoter of Example 4.2, but with the binding energies replaced by kinetic rate constants. ∇

Example 4.4 (Transcription of mRNA)**.** Consider the production of mRNA from a single copy of DNA. We have two basic reactions that can occur: mRNA can be produced by RNA polymerase transcribing the DNA and producing an mRNA strand, or mRNA can be degraded. We represent the microstate q of the system in terms of the number of mRNA's that are present, which we write as n for ease of notation. The reactions can now be represented as $\xi_1 = +1$, corresponding to transcription and $\xi_2 = -1$, corresponding to degradation. We choose as our propensity functions

$$a_1(n,t) = \alpha, \qquad a_2(n,t) = \delta n,$$

by which we mean that the probability of that a gene is transcribed in time dt is αdt and the probability that a transcript is created in time dt is $\delta n dt$ (proportional to the number of mRNA's).

We can now write down the master equation as described above. Equation (4.4) becomes

$$\begin{aligned} P(n,t+dt) &= P(n,t)\Big(1 - \sum_{i=1,2} a_i(n,t)dt\Big) + \sum_{i=1,2} P(n-\xi_i,t)a_i(q-\xi_i)dt \\ &= P(n,t) - a_1(n,t)P(n,t) - a_2(n,t)P(n,t) \\ &\qquad + a_1(n-1,t)P(n-1,t) + a_2(n+1,t)P(n+1) \\ &= P(n,t) + \alpha P(n-1,t)dt - (\alpha - \delta n)P(n,t)dt + \delta(n+1)P(n+1,t)dt. \end{aligned}$$

This formula holds for $n > 0$, with the $n = 0$ case satisfying

$$P(0,t+dt) = P(0,t) - \alpha P(0,t)dt + \delta P(1,t)dt.$$

Notice that we have an infinite number of equations, since n can be any positive integer.

We can write the differential equation version of the master equation by subtracting the first term on the right hand side and dividing by dt:

$$\frac{d}{dt}P(n,t) = \alpha P(n-1,t) - (\alpha+\delta n)P(n,t) + \delta(n+1)P(n+1,t), \qquad n>0$$
$$\frac{d}{dt}P(0,t) = -\alpha P(0,t)dt + \delta P(1,t).$$

Again, this is an infinite number of differential equations, although we could take some limit N and simply declare that $P(N,t) = 0$ to yield a finite number.

One simple type of analysis that can be done on this equation without truncating it to a finite number is to look for a steady state solution to the equation. In this case, we set $\dot{P}(n,t) = 0$ and look for a constant solution $P(n,t) = p_e(n)$. This yields an algebraic set of relations

$$\begin{aligned}
0 &= -\alpha p_e(0) + \delta p_e(1) && \Longrightarrow && \alpha p_e(0) &= \delta p_e(1)\\
0 &= \alpha p_e(0) - (\alpha+\delta)p_e(1) + 2\delta p_e(2) && && \alpha p_e(1) &= 2\delta p_e(2)\\
0 &= \alpha p_e(1) - (\alpha+2\delta)p_e(2) + 3\delta p_e(3) && && \alpha p_e(1) &= 3\delta p_e(3)\\
&\vdots && && &\vdots\\
& && && \alpha p(n-1) &= n\delta p(n).
\end{aligned}$$

It follows that the distribution of steady state probabilities is given by the Poisson distribution

$$p(n) = e^{\alpha/\delta}\frac{(\alpha/\delta)^n}{n!},$$

and the mean, variance and coefficient of variation are thus

$$\mu = \frac{\alpha}{\delta}, \qquad \sigma^2 = \frac{\alpha}{\delta}, \qquad CV = \frac{\mu}{\sigma} = \frac{1}{\sqrt{\mu}} = \sqrt{\frac{\delta}{\alpha}}.$$

Note taht the coefficient of variation increases if μ decreases. ∇

Chemical Langevin equation (CLE)

The chemical master equation gives a complete description of the evolution of the distribution of a system, but it can often be quite cumbersome to work with directly. A number of approximations to the master equation are thus used to provide more tractable formulations of the dynamics. The first of these that we shall consider is known as the *chemical Langevin equation* (CLE).

To derive the chemical Langevin equation, we start by assuming that the number of molecules in the system is large and that we can therefore represent the system using a vector of real numbers X, with X_i representing the (real-valued) number of molecules in S_i. (Often X_i will be divided by the volume to give a real-valued

concentration of species S_i.) In addition, we assume that we are interested in the dynamics on time scales in which individual reactions are not important and so we can look at how the system state changes over time intervals in which many reactions occur and hence the system state evolves in a smooth fashion.

Let $X(t)$ be the state vector for the system, where we assume now that the elements of X are real-valued rather than integer valued. We make the further approximation that we can lump together multiple reactions so that instead of keeping track of the individual reactions, we can average across a number of reactions over a time τ to allow the continuous state to evolve in continuous time. The resulting dynamics can be described by a stochastic process of the form

$$X_i(t+\tau) = X_i(t) + \sum_{j=1}^{M} \xi_{ij} a_j(X(t))\tau + \sum_{j=1}^{M} \xi_{ij} a_j^{1/2}(X(t)) \mathcal{N}_j(0, \sqrt{\tau}),$$

where a_j are the propensity functions for the individual reactions, ξ_{ij} are the corresponding changes in the system states X_i and $\mathcal{N}_j$ are a set of independent Gaussian random variables with zero mean and variance τ.

If we assume that τ is small enough that we can use the derivative to approximate the previous equation (but still large enough that we can average over multiple reactions), then we can write

$$\frac{dX_i(t)}{dt} = \sum_{j=1}^{M} \xi_{ji} a_j(X(t)) + \sum_{j=1}^{M} \xi_{ji} a_j^{1/2}(X(t)) \Gamma_j(t) \quad =: A_i(X(t)) + \sum_{j=1}^{M} B_{ij}(X(t)) \Gamma_j(t), \tag{4.9}$$

where Γ_j are white noise processes (see Appendix **??**). This equation is called the *chemical Langevin equation* (CLE).

Example 4.5 (Protein production)**.** Consider a simplified two-step model of protein production in which mRNA is produced by DNA and protein by mRNA. We do not model the detailed processes of isomerization and elongation of the mRNA and polypeptide chains. We can capture the state of the system by keeping track of the number of copies of DNA, mRNA, and protein, which we denote by X_D, X_m and X_P, respectively, so that $X = (X_D, X_m, X_P)$.

The simplified reactions with the corresponding propensity functions are given by

$$\begin{aligned}
&\mathrm{R}_1 : \mathrm{DNA} \xrightarrow{\alpha} \mathrm{mRNA} && \xi_1 = (1,0) && a_1(X) = \alpha\, X_D \\
&\mathrm{R}_2 : \mathrm{mRNA} \xrightarrow{\delta} \phi && \xi_2 = (-1,0) && a_2(X) = \delta\, X_m \\
&\mathrm{R}_3 : \mathrm{mRNA} \xrightarrow{\kappa} \mathrm{mRNA} + \mathrm{protein} && \xi_3 = (0,1) && a_3(X) = \kappa\, X_m \\
&\mathrm{R}_4 : \mathrm{protein} \xrightarrow{\gamma} \phi && \xi_4 = (0,-1) && a_4(X) = \gamma\, X_P.
\end{aligned}$$

Using these, we can write the Langevin equation as

$$\frac{dX_m}{dt} = \alpha X_D - \delta X_m + \sqrt{\alpha X_D}\Gamma_1(t) - \sqrt{\delta X_m}\Gamma_2(t)$$
$$\frac{dX_P}{dt} = \kappa X_m - \gamma X_P + \sqrt{\kappa X_m}\Gamma_3(t) - \sqrt{\gamma X_p}\Gamma_4(t).$$

We can keep track of the species concentration by dividing the number of molecules by the volume Ω. Letting $m = X_m/\Omega$, $P = X_P/\Omega$, and $\alpha_0 = \alpha X_D/\Omega$, we obtain the final expression

$$\frac{d}{dt}\begin{pmatrix} m \\ P \end{pmatrix} = \begin{pmatrix} -\delta & 0 \\ \kappa & -\gamma \end{pmatrix}\begin{pmatrix} m \\ P \end{pmatrix} + \begin{pmatrix} \alpha_0 \\ 0 \end{pmatrix} + \frac{1}{\sqrt{\Omega}}\begin{pmatrix} \left(\sqrt{\alpha_0 + \delta m}\right)\Gamma_m \\ \left(\sqrt{\kappa m + \gamma P}\right)\Gamma_P \end{pmatrix},$$

where Γ_m and Γ_P are independent white noise processes with unit variance (note that here we have used that if Γ_1 and Γ_2 are Gaussian white noises with unit variance, then $\sqrt{a}\Gamma_1 + \sqrt{b}\Gamma_2 = \sqrt{a+b}\Gamma$ with Γ also a Gaussian white noise with unit variance). ∇

Fokker-Planck equations (FPE)

The chemical Langevin equation provides a stochastic ordinary differential equation that describes the evolution of the system state. A slightly different (but completely equivalent) representation of the dynamics is to model how the probability distribution $P(x,t)$ evolves in time. As in the case of the chemical Langevin equation, we will assume that the system state is continuous and write down a formula for the evolution of the density function $p(x,t)$. This formula is known as the *Fokker-Planck equations* (FPE) and is essentially an approximation on the chemical master equation.

Consider first the case of a random process in one dimension. We assume that the random process is in the same form as the previous section:

$$\frac{dX(t)}{dt} = A(X(t)) + B(X(t))\Gamma(t). \tag{4.10}$$

The function $A(X)$ is called the *drift term* and $B(X)$ is the *diffusion term*. It can be shown that the probability density function for X, $p(x,t)$, satisfies the partial differential equation

$$\frac{\partial p}{\partial t}(x,t) = -\frac{\partial}{\partial x}(A(x,t)p(x,t)) + \frac{1}{2}\frac{\partial^2}{\partial x^2}(B^2(x,t)p(x,t)) \tag{4.11}$$

Note that here we have shifted to the probability density function since we are considering X to be a continuous state random process.

In the multivariate case, a bit more care is required. Using the chemical Langevin equation (4.9), we define

$$D_i(x,t) = \sum_{j=1}^{M} B_{ij}^2(x,t), \qquad C_{ij}(x,t) = \sum_{k=1}^{M} B_{ik}(x,t)B_{jk}(x,t),\ i < j = 1,\ldots,M.$$

The Fokker-Planck equation now becomes

$$\frac{\partial p}{\partial t}(x,t) = -\sum_{i=1}^{M} \frac{\partial}{\partial x_i}(A_i(x,t)p(x,t)) + \frac{1}{2}\sum_{i=1}^{M} \frac{\partial}{\partial x_i}\frac{\partial^2}{\partial x^2}(D_i(x,t)p(x,t)) + \sum_{\substack{i,j=1 \\ i<j}}^{M} \frac{\partial^2}{\partial x_i \partial x_j}(C_{ij}(x,t)p(x,t)). \quad (4.12)$$

Note that the Fokker-Planck equation is very similar to the chemical master equation: both provide a description of how the probability distribution varies as a function of time. In the case of the Fokker-Planck equation, we regard the state as a continuous set of variables and we write a partial differential equation for how the probability density function evolves in time. In the case of the chemical master equation, we have a discrete state (microstates) and we write an ordinary differential equation for how the probability distribution (formally the probability mass function) evolves in time. Both formulations contain the same basic information, just using slightly different representations of the system and the probability of being in a given state.

Linear noise approximation (LNA)

The chemical Langevin equation and the Fokker-Planck equation provide approximations to the chemical master equation. A slightly different approximation can be obtained by expanding the density function in terms of a size parameter Ω. This approximation is know as the *linear noise approximation* (LNA) or the Ω *expansion* [49].

We begin with the master equation for a *continuous* random variable X. Formally deriving this requires a considerable effort since we have to extend our previous discussions to the case where the random variable has a continuous set of values. To do this, we rewrite the propensity function $a_i(q,t)$ as $a_\xi(q,t;\Omega)$, where $q \in \mathbb{R}^n$ is a vector of continuous states and $\xi \in \mathbb{R}^n$ is a vector of continuous "increments" (the analog of reactions). We also explicitly keep track of the dependence of the propensity function on a parameter Ω (the volume in our case).

Using this notation, we can write the master equation for the random variable X as

$$\frac{\partial P}{\partial t}(x,t) = \int (a_\xi(x-\xi,t;\Omega)P(x-\xi,t) - a_\xi(x,t;\Omega)P(x,t))\,d\xi.$$

Since we are working with continuous variables, we now have an integral in place of our previous sum. In addition, if we take the derivative of $P(x,t)$ with respect to the continuous variable x, we can obtain the pdf of the distribution $p(x,t)$ and this satisfies the equation

$$\frac{\partial p}{\partial t}(x,t) = \int (a_\xi(x-\xi,t;\Omega)p(x-\xi,t) - a_\xi(x,t;\Omega)p(x,t))\,d\xi.$$

Although we are skipping important theoretical details, the basic idea of this formulation is the same as the discrete chemical master equation: we keep track of how the probability density changes by "summing" (integrating) over all (incremental) reactions going into and out of that particular state.

We now assume that the mean of X can be written as $\Omega\phi(t)$ where $\phi(t)$ is a continuous function of time that represents the evolution of the mean of X/Ω. To understand the fluctuations of the system about this mean, we write

$$X = \Omega\phi + \Omega^{\frac{1}{2}}Z,$$

where Z is a new variable representing the perturbations of the system about its mean. We can write the distribution for Z as

$$p_Z(z,t) = p_X(\Omega\phi(t) + \Omega^{\frac{1}{2}}z, t)$$

and it follows that the derivatives of p_Z can be written as

$$\frac{\partial^\nu p_Z}{z^\nu} = \Omega^{\frac{1}{2}\nu}\frac{\partial^\nu p_X}{x^\nu}$$
$$\frac{\partial p_Z}{\partial t} = \frac{\partial p_X}{\partial t} + \Omega\frac{d\phi}{dt}\frac{\partial p_X}{\partial x} = \frac{\partial p_X}{\partial t} + \Omega^{\frac{1}{2}}\frac{d\phi}{dt}\frac{\partial p_Z}{\partial z}.$$

We further assume that the Ω dependence of the propensity function is such that

$$a_\xi(\Omega\phi,t;\Omega) = f(\Omega)\tilde{a}_\xi(\phi),$$

where $\tilde{a}$ is not dependent on the parameter Ω or the time t. From these relations, we can now derive the master equation for p_Z in terms of powers of Ω (derivation omitted).

The $\Omega^{1/2}$ term in the expansion turns out to yield

$$\frac{d\phi}{dt} = \int \xi\, a_\xi(\Omega\phi)\,d\xi, \qquad \phi(0) = \frac{X(0)}{\Omega},$$

which is precisely the equation for the mean of the concentration. It can further be shown that the terms in Ω^0 are given by

$$\frac{\partial p_Z(z,\tau)}{\partial\tau} = -\alpha_1'(\phi)\frac{\partial}{\partial z}(z p_Z(z,t)) + \frac{1}{2}\alpha_2(\phi)\frac{\partial^2 p_Z(z,t)}{\partial z^2}, \tag{4.13}$$

where

$$\alpha_\nu(x) = \int \xi^\nu \tilde{a}_\xi(x)\, d\xi, \qquad \tau = \Omega^{-1} f(\Omega) t.$$

Notice that in the case that $\phi(t) = \phi_0$ (a constant), this equation becomes the Fokker-Planck equation derived previously.

Higher order approximations to this equation can also be carried out by keeping track of the expansion terms in higher order powers of Ω. In the case where Ω represents the volume of the system, the next term in the expansion is Ω^{-1} and this represents fluctuations that are on the order of a single molecule, which can usually be ignored.

Reaction rate equations (RRE)

As we already saw in Chapter 2, the reaction rate equations can be used to describe the dynamics of a chemical system in the case where there are a large number of molecules whose state can be approximated using just the concentrations of the molecules. We re-derive the results from Section 2.1 here, being more careful to point out what approximations are being made.

We start with the chemical Langevin equations (4.9), from which we can write the dynamics for the average quantity of the each species at each point in time:

$$\frac{d\langle X_i(t)\rangle}{dt} = \sum_{j=1}^{M} \xi_{ji} \langle a_j(X(t))\rangle, \tag{4.14}$$

where the second order term drops out under the assumption that the Γ_j's are independent processes with zero mean. We see that the reaction rate equations follow by defining $x_i = \langle X_i\rangle/\Omega$ and *assuming* that $\langle a_j(X(t))\rangle = a_j(\langle X(t)\rangle)$. This relationship is true when a_j is linear (e.g., in the case of a unimolecular reaction), but is an approximation otherwise.

4.2 Simulation of Stochastic Systems

Suppose that we want to generate a collection of sample trajectories for a stochastic system whose evolution is described by the chemical master equation (4.5):

$$\frac{d}{dt}P(q,t) = \sum_i a_i(q-\xi_i)P(q-\xi_i,t) - \sum_i a_i(q)P(q,t),$$

where $P(q,t)$ is the probability of being in a microstate q at time t (starting from q_0 at time t_0) and $a_i(q)$ is the propensity function for a reaction i starting at a microstate q and ending at microstate $q+\xi_i$. Instead of simulating the distribution function $P(q,t)$, we wish to simulate a specific instance $q(t)$ starting from some initial condition $q_0(t_0)$. If we simulate many such instances of $q(t)$, their distribution at time t should match $P(q,t)$.

A naive algorithm for stochastic simulation

The Stochastic Simuilation Algorithm

To illustrate the basic ideas that we will use, consider first a simple birth process in which the microstate is given by an integer $q \in \{0, 1, 2, \dots\}$ and we assume that the propensity function is given by

$$a(q)\,dt = \lambda\,dt, \qquad \xi = +1.$$

Thus the probability of transition increases linearly with the time increment dt (so birth events occur at rate λ, on average). If we assume that the birth events are independent of each other, then it can be shown (see Appendix **??**) that this process has Poisson distribution with parameter $\lambda\tau$:

$$P(q(t+\tau) - q(t) = \ell) = \frac{(\lambda\tau)^\ell}{\ell!} e^{-\lambda\tau},$$

where τ is the difference in time and ℓ is the difference in count q. In fact, this distribution is a joint distribution in time τ and count ℓ, and by setting $\ell = 1$ it can be seen that the time to the next reaction T follows an exponential distribution and has density function

$$p_T(\tau) = \lambda\, e^{-\lambda\tau}.$$

The exponential distribution has expectation $1/\lambda$ and so we see that the average time between events is inversely proportional to the reaction rate λ.

Consider next a more general case in which we have a countable number of microstates $q \in \{0, 1, 2, \dots\}$ and we let k_{ji} represent the transition probability between a microstate i and microstate j. The birth process is a special case given by $k_{i+1,i} = \lambda$ and all other $k_{ji} = 0$. The chemical master equation describes the joint probability that we are in state $q = i$ at a particular time t. We would like to know the probability that we transition to a new state $q = j$ at time $t + dt$. Given this probability, we can attempt to generate an instance of the variable $q(t)$ by first determining which reaction occurs and then when the reaction occurs.

Let $P(j, \tau) := P(j, t+\tau+d\tau \mid i, t+\tau)$ represent the probability that we transition from the state i to the state j in the time interval $[t+\tau, t+\tau+d\tau]$. For simplicity and ease of notation, we will take $t = 0$. Let $T := T_{j,i}$ be the time at which the reaction first occurs. We can write the probability that we transition to state j in the interval $[\tau, \tau + d\tau]$ as

$$P(j, \tau) = P(T > \tau)\, k_{ji}\, d\tau, \tag{4.15}$$

where $P(T > \tau)$ is the probability that no reaction occurs in the time interval $[0, \tau]$ and $k_{ji} d\tau$ is the probability that the reaction taking state i to state j occurs in the next $d\tau$ seconds (assumed to be independent events, giving the product of these probabilities).

To compute $P(T > \tau)$, define

$$\bar{k}_i = \sum_j k_{ji}$$

so that $(1 - \bar{k}_i)d\tau$ is the probability that no transition occurs from state i in the next $d\tau$ seconds. Then, the probability that no reaction occurs in the interval $[\tau, \tau + d\tau]$ can be written as

$$P(T > \tau + d\tau) = P(T > \tau)(1 - \bar{k}_i)\, d\tau. \tag{4.16}$$

It follows that

$$\frac{d}{d\tau}P(T > \tau) = \lim_{d\tau \to 0} \frac{P(T > \tau + d\tau) - P(T > \tau)}{d\tau} = -P(T > \tau)\, \bar{k}_i.$$

Solving this differential equation, we obtain

$$P(T > \tau) = e^{-\bar{k}_i \tau}, \tag{4.17}$$

so that the probability that no reaction occurs in time τ decreases exponentially with the amount of time that we wait, with rate given by the sum of all the reactions that can occur from state i.

We can now combine equation (4.17) with equation (4.15) to obtain

$$P(j, \tau) = P(j, \tau + d\tau \mid i, 0) = k_{ji} e^{-\bar{k}_i \tau}\, d\tau.$$

We see that this has the form of a density function in time and hence the probability that the next reaction is reaction j, independent of the time in which it occurs, is

$$P_{ji} = \int_0^\infty k_{ji} e^{-\bar{k}_i \tau}\, d\tau = \frac{k_{ji}}{\bar{k}_i}. \tag{4.18}$$

Thus, to choose the next reaction to occur from a state i, we choose between N possible reactions, with the probability of each reaction weighted by $k_{ji}/\bar{k}_i$.

To determine the time that the next reaction occurs, we sum over all possible reactions j to get the density function for the reaction time:

$$p_T(\tau) = \sum_j k_{ji} e^{-\bar{k}_i \tau} = \bar{k}_i e^{-\bar{k}_i \tau}.$$

This is the density function associated with a Poisson distribution. To compute a time of reaction Δt that draws from this distribution, we note that the cumulative distribution function for T is given by

$$\int_0^{\Delta t} f_T(\tau)\, d\tau = \int_0^{\Delta t} \bar{k}_i e^{-\bar{k}_i \tau}\, d\tau = 1 - e^{-\bar{k}_i \Delta t}.$$

The cumulative distribution function is always in the range $[0,1]$ and hence we can compute Δt by choosing a (uniformly distributed) random number r in $[0,1]$ and then computing

$$\Delta t = \frac{1}{k_i} \ln \frac{1}{1-r}. \tag{4.19}$$

(This equation can be simplified somewhat by replacing $1-r$ with r' and noting that r' can also be drawn from a uniform distribution on $[0,1]$.)

Note that in the case of a birth process, this computation agrees with our earlier analysis. Namely, $\overline{k_i} = \lambda$ and hence the (only) reaction occurs according to an exponential distribution with parameter λ.

This set of calculations gives the following algorithm for computing an instance of the chemical master equation:

1. Choose an initial condition q at time $t = 0$.
2. Calculate the propensity functions $a_i(q)$ for each possible reaction q.
3. Choose the time for the reaction according to equation (4.19), where $r \in [0,1]$ is chosen from a uniform distribution.
4. Use a weighted random number generator to identify which reaction will take place next, using the weights in equation (4.18).
5. Update q by implementing the reaction ξ and update the time t by δt
6. If $T < T_{\text{stop}}$, goto step 2.

This method is sometimes called "Gillespie's direct method" [29, 30], but we shall refer to it here as the "stochastic simulation algorithm" (SSA). We note that the reaction number in step 4 can be computed by calculating a uniform random number on $[0,1]$, scaling this by the total propensity $\sum_i a_i(\xi_i, q)$, and then finding the first reaction i such that $\sum_{j=0}^{i} a(\xi_j, q)$ is larger than this scaled random number.

Example 4.6 (Transcription)**.** To be completed. ∇ Review

4.3 Input/Output Linear Stochastic Systems

In many situations, we wish to noise how noise propogates through a biomolecular system. For example, we may wish to understand how stochastic variations in RNA polymerase concentraton affect gene expression. In order to analyze these cases, we specialize to the case of a biomolecular system operating around a fixed operating point.

We now consider the problem of how to compute the response of a linear system to a random process. We assume we have a linear system described in state space as

$$\dot{X} = AX + FW, \qquad Y = CX \tag{4.20}$$

Given an "input" W, which is itself a random process with mean $\mu(t)$, variance $\sigma^2(t)$ and correlation $r(t, t+\tau)$, what is the description of the random process Y?

Let W be a white noise process, with zero mean and noise intensity Q:

$$r(\tau) = Q\delta(\tau).$$

We can write the output of the system in terms of the convolution integral

$$Y(t) = \int_0^t h(t-\tau)W(\tau)\,d\tau,$$

where $h(t-\tau)$ is the impulse response for the system

$$h(t-\tau) = Ce^{A(t-\tau)}B + D\delta(t-\tau).$$

We now compute the statistics of the output, starting with the mean:

$$\begin{aligned}\mathbb{E}(Y(t)) &= \mathbb{E}\Big(\int_0^t h(t-\eta)W(\eta)\,d\eta\Big)\\ &= \int_0^t h(t-\eta)\mathbb{E}(W(\eta))\,d\eta = 0.\end{aligned}$$

Note here that we have relied on the linearity of the convolution integral to pull the expectation inside the integral.

We can compute the covariance of the output by computing the correlation $r_Y(\tau)$ and setting $\sigma_Y^2 = r_Y(0)$. The correlation function for y is

$$\begin{aligned}r_Y(t,s) = \mathbb{E}(Y(t)Y(s)) &= \mathbb{E}\Big(\int_0^t h(t-\eta)W(\eta)\,d\eta \cdot \int_0^s h(s-\xi)W(\xi)\,d\xi\Big)\\ &= \mathbb{E}\Big(\int_0^t\int_0^s h(t-\eta)W(\eta)W(\xi)h(s-\xi)\,d\eta d\xi\Big)\end{aligned}$$

Once again linearity allows us to exchange expectation and integration

$$\begin{aligned}r_Y(t,s) &= \int_0^t\int_0^s h(t-\eta)\mathbb{E}(W(\eta)W(\xi))h(s-\xi)\,d\eta d\xi\\ &= \int_0^t\int_0^s h(t-\eta)Q\delta(\eta-\xi)h(s-\xi)\,d\eta d\xi\\ &= \int_0^t h(t-\eta)Qh(s-\eta)\,d\eta\end{aligned}$$

Now let $\tau = s-t$ and write

$$\begin{aligned}r_Y(\tau) = r_Y(t,t+\tau) &= \int_0^t h(t-\eta)Qh(t+\tau-\eta)\,d\eta\\ &= \int_0^t h(\xi)Qh(\xi+\tau)\,d\xi \qquad \text{(setting } \xi = t-\eta\text{)}\end{aligned}$$

Finally, we let $t \to \infty$ (steady state)

$$\lim_{t\to\infty} r_Y(t, t+\tau) = \overline{r}_Y(\tau) = \int_0^\infty h(\xi)Qh(\xi+\tau)d\xi \tag{4.21}$$

If this integral exists, then we can compute the second order statistics for the output Y.

We can provide a more explicit formula for the correlation function r in terms of the matrices A, F and C by expanding equation (4.21). We will consider the general case where $W \in \mathbb{R}^p$ and $Y \in \mathbb{R}^q$ and use the correlation matrix $R(t,s)$ instead of the correlation function $r(t,s)$. Define the *state transition matrix* $\Phi(t,t_0) = e^{A(t-t_0)}$ so that the solution of system (4.20) is given by

$$x(t) = \Phi(t,t_0)x(t_0) + \int_{t_0}^t \Phi(t,\lambda)Fw(\lambda)d\lambda$$

Proposition 4.1 (Stochastic response to white noise)**.** *Let $\mathbb{E}(X(t_0)X^T(t_0)) = P(t_0)$ and W be white noise with $\mathbb{E}(W(\lambda)W^T(\xi)) = R_W\delta(\lambda-\xi)$. Then the correlation matrix for X is given by*

$$R_X(t,s) = P(t)\Phi^T(s,t)$$

where $P(t)$ satisfies the linear matrix differential equation

$$\dot{P}(t) = AP + PA^T + FR_WF, \qquad P(0) \quad = P_0.$$

Proof. Using the definition of the correlation matrix, we have

$$\begin{aligned}
\mathbb{E}(X(t)X^T(s)) &= \mathbb{E}\Big(\Phi(t,0)X(0)X^T(0)\Phi^T(t,0) + \text{cross terms} \\
&\qquad + \int_0^t \Phi(t,\xi)FW(\xi)\,d\xi \int_0^s W^t(\lambda)F^T\Phi(s,\lambda)\,d\lambda\Big) \\
&= \Phi(t,0)\mathbb{E}(X(0)X^T(0))\Phi(s,0) \\
&\qquad + \int_0^t\int_0^s \Phi(t,\xi)F\mathbb{E}(W(\xi)W^T(\lambda))F^T\Phi(s,\lambda)\,d\xi\,d\lambda \\
&= \Phi(t,0)P(0)\phi^T(s,0) + \int_0^t \Phi(t,\lambda)FR_W(\lambda)F^T\Phi(s,\lambda)\,d\lambda.
\end{aligned}$$

Now use the fact that $\Phi(s,0) = \Phi(s,t)\Phi(t,0)$ (and similar relations) to obtain

$$R_X(t,s) = P(t)\Phi^T(s,t)$$

where

$$P(t) = \Phi(t,0)P(0)\Phi^T(t,0) + \int_0^T \Phi(t,\lambda)FR_WF^T(\lambda)\Phi^T(t,\lambda)d\lambda$$

Finally, differentiate to obtain

$$\dot{P}(t) = AP + PA^T + FR_WF, \qquad P(0) \quad = P_0$$

(see Friedland [26] for details). □

The correlation matrix for the output Y can be computed using the fact that $Y = CX$ and hence $R_Y = C^T R_X C$. We will often be interested in the steady state properties of the output, which are given by the following proposition.

Proposition 4.2 (Steady state response to white noise)**.** *For a time-invariant linear system driven by white noise, the correlation matrices for the state and output converge in steady state to*

$$R_X(\tau) = R_X(t, t+\tau) = Pe^{A^T\tau}, \qquad R_Y(\tau) = CR_X(\tau)C^T$$

where P satisfies the algebraic equation

$$AP + PA^T + FR_W F^T = 0 \qquad P > 0. \tag{4.22}$$

Equation (4.22) is called the *Lyapunov equation* and can be solved in MATLAB using the function `lyap`.

Example 4.7 (First-order system)**.** Consider a scalar linear process

$$\dot{X} = -aX + W, \qquad Y = cX,$$

where W is a white, Gaussian random process with noise intensity σ^2. Using the results of Proposition 4.1, the correlation function for X is given by

$$R_X(t, t+\tau) = p(t)e^{-a\tau}$$

where $p(t) > 0$ satisfies

$$p(t) = -2ap + \sigma^2.$$

We can solve explicitly for $p(t)$ since it is a (non-homogeneous) linear differential equation:

$$p(t) = e^{-2at}p(0) + (1 - e^{-2at})\frac{\sigma^2}{2a}.$$

Finally, making use of the fact that $Y = cX$ we have

$$r(t, t+\tau) = c^2(e^{-2at}p(0) + (1 - e^{-2at})\frac{\sigma^2}{2a})e^{-a\tau}.$$

In steady state, the correlation function for the output becomes

$$r(\tau) = \frac{c^2\sigma^2}{2a}e^{-a\tau}.$$

Note correlation function has the same form as the Ornstein-Uhlenbeck process in Example **??** (with $Q = c^2\sigma^2$). ∇

As in the case of deterministic linear systems, we can analyze a stochastic linear system either in the state space or the frequency domain. The frequency domain approach provides a very rich set of tools for modeling and analysis of interconnected systems, relying on the frequency response and transfer functions to represent the flow of signals around the system.

Given a random process $X(t)$, we can look at the frequency content of the properties of the response. In particular, if we let $\rho(\tau)$ be the correlation function for a (scalar) random process, then we define the *power spectral density function* as the Fourier transform of ρ:

$$S(\omega) = \int_{-\infty}^{\infty} \rho(\tau) e^{-j\omega\tau}\, d\tau, \qquad \rho(\tau) \quad = \frac{1}{2\pi} \int_{-\infty}^{\infty} S(\omega) e^{j\omega\tau}\, d\tau.$$

The power spectral density provides an indication of how quickly the values of a random process can change through the frequency content: if there is high frequency content in the power spectral density, the values of the random variable can change quickly in time.

Example 4.8 (Ornstein-Uhlenbeck process)**.** To illustrate the use of these measures, consider a first-order Markov process where the correlation function is

$$\rho(\tau) = \frac{Q}{2\omega_0} e^{-\omega_0(\tau)}.$$

This correspnds to Example 4.7 (also called an *Ornstein-Uhlenbeck process*). The power spectral density becomes

$$\begin{aligned} S(\omega) &= \int_{-\infty}^{\infty} \frac{Q}{2\omega_0} e^{-\omega|\tau|} e^{-j\omega\tau}\, d\tau \\ &= \int_{-\infty}^{0} \frac{Q}{2\omega_0} e^{(\omega - j\omega)\tau}\, d\tau + \int_{0}^{\infty} \frac{Q}{2\omega_0} e^{(-\omega - j\omega)\tau}\, d\tau = \frac{Q}{\omega^2 + \omega_0^2}. \end{aligned}$$

We see that the power spectral density is similar to a transfer function and we can plot $S(\omega)$ as a function of ω in a manner similar to a Bode plot, as shown in Figure 4.2. Note that although $S(\omega)$ has a form similar to a transfer function, it is a real-valued function and is not defined for complex s. ∇

Using the power spectral density, we can more formally define "white noise": a *white noise process* is a zero-mean, random process with power spectral density $S(\omega) = W$ = constant for all ω. If $X(t) \in \mathbb{R}^n$ (a random vector), then $W \in \mathbb{R}^{n\times n}$. We see that a random process is white if all frequencies are equally represented in its power spectral density; this spectral property is the reason for the terminology "white".

Given a linear system

$$\dot{X} = AX + FW, \qquad Y = CX,$$

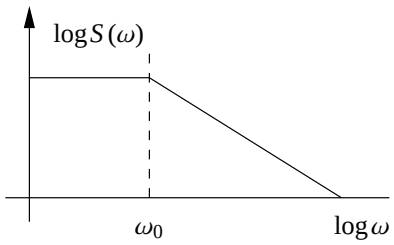


Figure 4.2: Spectral power density for a first-order Markov process.

.

with W given by white noise, we can compute the spectral density function corresponding to the output Y. We start by computing the Fourier transform of the steady state correlation function (4.21):

$$\begin{aligned} S_Y(\omega) &= \int_{-\infty}^{\infty} \left[\int_0^{\infty} h(\xi) Q h(\xi+\tau) d\xi \right] e^{-j\omega\tau} d\tau \\ &= \int_0^{\infty} h(\xi) Q \left[\int_{-\infty}^{\infty} h(\xi+\tau) e^{-j\omega\tau} d\tau \right] d\xi \\ &= \int_0^{\infty} h(\xi) Q \left[\int_0^{\infty} h(\lambda) e^{-j\omega(\lambda-\xi)} d\lambda \right] d\xi \\ &= \int_0^{\infty} h(\xi) e^{j\omega\xi} d\xi \cdot QH(j\omega) = H(-j\omega) Q H(j\omega). \end{aligned}$$

This is then the (steady state) response of a linear system to white noise.

As with transfer functions, one of the advantages of computations in the frequency domain is that the composition of two linear systems can be represented by multiplication. In the case of the power spectral density, if we pass white noise through a system with transfer function $H_1(s)$ followed by transfer function $H_2(s)$, the resulting power spectral density of the output is given by

$$S_Y(\omega) = H_1(-j\omega) H_2(-j\omega) Q_u H_2(j\omega) H_1(j\omega).$$

As stated earlier, white noise is an idealized signal that is not seen in practice. One of the ways to produced more realistic models of noise and disturbances is to apply a filter to white noise that matches a measured power spectral density function. Thus, we wish to find a covariance W and filter $H(s)$ such that we match the statistics $S(\omega)$ of a measured noise or disturbance signal. In other words, given $S(\omega)$, find $W > 0$ and $H(s)$ such that $S(\omega) = H(-j\omega) W H(j\omega)$. This problem is know as the *spectral factorization problem*.

Figure 4.3 summarizes the relationship between the time and frequency domains.

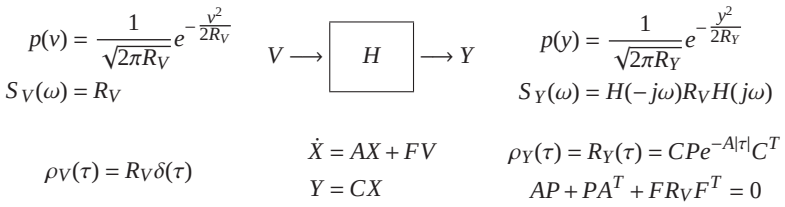


Figure 4.3: Summary of steady state stochastic response.

Exercises

4.1 (BE 150, Winter 2011) For this problem, we return to our standard model of transcription and transcription process with probabilistic creation and degradation of discrete mRNA and protein molecules. The *propensity functions* for each reaction are as follows:
Probability of transcribing 1 mRNA molecule: $0.2dt$
Probability of degrading 1 mRNA molecule: $0.5dt$ and is proportional to the number of mRNA molecules.
Probability of translating 1 protein: $5dt$ and is proportional to the number of mRNA molecules.
Probability of degrading 1 protein molecule: $0.5dt$ and is proportional to the number of protein molecules.
dt is the time step chosen for your simulation. Here we choose $dt = 0.05$.

(a) Simulate the stochastic system above until time $T = 100$. Plot the resulting number of mRNA and protein over time.

(b) Now assume that the proteins are degraded much more slowly than mRNA and the propensity function of protein degradation is now $0.05dt$. To maintain similar protein levels, the translation probability is now $0.5dt$ (and still proportional to the number of mRNA molecules). Simulate this system as above. What difference do you see in protein level? Comment on the effect of protein degradation rates on noise.

4.2 (BE 150, Winter 2011) Compare a simple model of negative autoregulation with one without autoregulation:

$$\frac{dX}{dt} = \beta_0 - \gamma X$$

and

$$\frac{dX}{dt} = \frac{\beta}{1 + \frac{X}{K}} - \gamma X$$

(a) Assume that the basal transcription rates β and β_0 vary between cells, following a Gaussian distribution with $\frac{\sigma^2}{<X>} = 0.1$. Simulate time courses of both models for

100 different "cells" using the following parameters: $\beta = 2, \beta_0 = 1, \gamma = 1, K = 1$. Plot the nonregulated and autoregulated systems in two separate plots. Comment on the variation you see in the time courses.

(b) Calculate the deterministic steady state for both models above. How does variation in the basal transcription rate β or β_0 enter into the steady state and relate it to what you see in part (a).

4.3 Consider gene expression: $\phi \xrightarrow{\alpha} \mathrm{m}$, $\mathrm{m} \xrightarrow{\kappa} \mathrm{m} + \mathrm{P}$, $\mathrm{m} \xrightarrow{\delta} \phi$, and $\mathrm{P} \xrightarrow{\gamma} \emptyset$. Answer the following questions:

(a) Use the stochastic simulation algorithm (SSA) to obtain realizations of the stochastic process of gene expression and numerically compare with the deterministic ODE solution. Explore how the realizations become close to or apart from the ODE solution when the volume is changed. Determine the stationary probability distribution for the protein (you can do this numerically, but note that this process is linear, so you can compute the probability distribution analytically in closed form).

(b) Now consider the additional binding reaction of protein P with downstream DNA binding sites D: $\mathrm{P} + \mathrm{D} \underset{k_{off}}{\overset{k_{on}}{\rightleftharpoons}} \mathrm{C}$. Note that the system no longer linear due to the presence of a bi-molecular reaction. Use the SSA algorithm to obtain sample realizations and numerically compute the probability distribution of the protein and compare it to what you obtained in part (a). Explore how this probability distribution and the one of C change as the rates k_{on} and k_{off} become larger and larger with respect to $\gamma, \alpha, \kappa, \delta$. Do you think we can use a QSS approximation similar to what we have done for ODE models?

(c) Determine the Langevin equation for the system in part (b) and obtain sample realizations. Explore numerically how good this approximation is when the volume decreases/increases.

4.4 Consider the bi-molecular reaction $\mathrm{A} + \mathrm{B} \underset{k_2}{\overset{k_1}{\rightleftharpoons}} \mathrm{C}$, in which A and B are in total amounts A_T and B_T, respectively. Compare the steady state value of C obtained from the deterministic model to the mean value of C obtained from the stochastic model as the volume is changed in the stochastic model. What do you observe? You can perform this investigation through numerical simulation.

4.5 Consider the simple birth and death process: $\mathrm{Z} \underset{k_1 G}{\overset{k_2 G}{\rightleftharpoons}} \emptyset$, in which G is a "gain". Assume that the reactions are catalyzed by enzymes and that the gain G can be tuned by changing the amounts of these enzymes. A deterministic ODE model for this system incorporating noise and disturbances due to the stochasticity of the cellular environment is given by

$$\dot{Z} = k_1 G - k_2 G Z + d(t),$$

in which $d(t)$ incorporates noise, as seen in the previous homework. Determine the Langevin equation for this birth and death process and compare its form to the deterministic one. Also, determine the frequency response of Z to noise for both the deterministic model and for the Langevin model. Does increasing the gain G has the same effect in both models? Explain.

4.6 Consider a second order system with dynamics

$$\begin{pmatrix} \dot{X}_1 \\ \dot{X}_2 \end{pmatrix} = \begin{pmatrix} -a & 0 \\ 0 & -b \end{pmatrix} \begin{pmatrix} X_1 \\ X_2 \end{pmatrix} + \begin{pmatrix} 1 \\ 1 \end{pmatrix} v, \qquad Y = \begin{pmatrix} 1 & 1 \end{pmatrix} \begin{pmatrix} X_1 \\ X_2 \end{pmatrix}$$

that is forced by Gaussian white noise with zero mean and variance σ^2. Assume $a, b > 0$.

(a) Compute the correlation function $\rho(\tau)$ for the output of the system. Your answer should be an explicit formula in terms of a, b and σ.

(b) Assuming that the input transients have died out, compute the mean and variance of the output.

Chapter 5
Feedback Examples

In this chapter we present a collection of examples that illustrate some of the modeling and analysis tools covered in the preceding chapters. Each of these examples represents a more complicated system than we have considered previous and together they are intended to demonstrate both the role of feedback in biological systems and how tools from control and dynamical systems can be applied to provide insight and understanding. Each of the sections below is indepedent of the others and they can be read in any order (or skipped entirely).

Pagination in this chapter is broken down by section to faciliate author editing. Review Some extraneous blank pages may be included due to LaTeX processing.

5.1 The *lac* Operon

The *lac* operon is one of the most studied regulatory networks in molecular biology. Its function is to determine when the cell should produce the proteins and enzymes necessary to import and metabolize lactose from its external environment. Since glucose is a more efficient source of carbon, the lactose machinery is not produced unless lactose is present and glucose is not present. The *lac* control system implements this computation.

In this section we construct a model for the lac operon and use that model to understand how changes of behavior can occur for large changes in parameters (e.g., lactose/glucose concentrations) and also the sensitivity of the phenotypic response to changes in individual parameter values in the model. The basic model and much of the analysis in this section is drawn from the work of Yildirim and Mackey [99].

Modeling

In constructing a model for the *lac* system, we need to decide what questions we wish to answer. Here we will attempt to develop a model that allows us to understand what levels of lactose are required for the *lac* system to become active in the absence of glucose. We will focus on the so-called "bistability" of the *lac* operon: there are two steady operating conditions—at low lactose levels the machinery is off and at high lactose levels the machinery is on. The system has hysteresis, so once the operon is actived, it remains active even if the lactose concentration descreases. We will construct a differential equation model of the system, with various simplifying assumptions along the way.

A schematic diagram of the *lac* control system is shown in Figure 5.1. Starting at the bottom of the figure, lactose permease is an integral membrane protein that helps transport lactose into the cell. Once in the cell, lactose is converted to allolactose, and allolactose is then broken down into glucose and galactose, both with the assistance of the enzyme β-galactosidase (β-gal for short). From here, the glucose is processed using the usual glucose metabolic pathway and the galactose.

The control circuitry is implemented via the reactions and transcriptional regulation shown in the top portion of the diagram. The *lac* operon, consisting of the genes *lacZ* (coding for β-gal), *lacY* (coding for lactose permease) and *lacA* (coding for a transacetylase), has a combinatorial promoter. Normally, lac repressor (*lacI*) is present and the operon is off. The activator for the operon is CAP, which has a positive inducer cAMP. The concentration of cAMP is controlled by glucose: when glucose is present, there is very little cAMP available in the cell (and hence CAP is not active).

The bistable switching behavior in the *lac* control system is implemented with a feedback circuit involving the *lac* repressor. Allolactose binds *lac* repressor and so when lactose is being metabolized, then the repressor is sequestered by allolactose and the *lac* operon is no longer repressed.

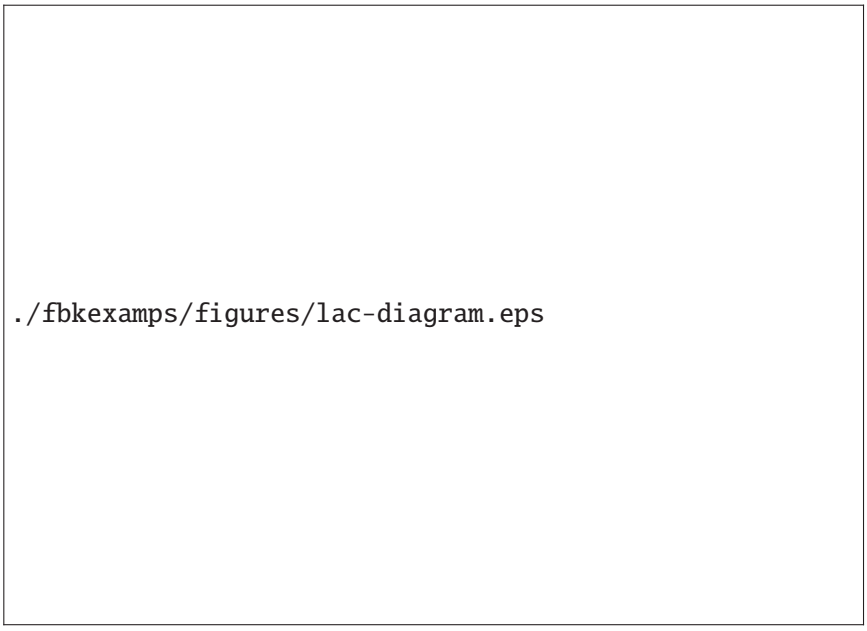


Figure 5.1: Schematic diagram for the *lac* system [99]. Permission pending.

To model this circuit, we need to write down the dynamics of all of the reactions and protein production. We will denote the concentration of the β-gal mRNA and protein as m_b and B. We assume that the internal concentration of lactose is given by L, ignoring the dynamics of lactose permease and transport of lactose into the cell. Similarly, we assume that the concentration of repressor protein, denoted R, is constant.

We start by keeping track of the concentration of free allolactose A. The relevant reactions are given by the transport of lactose into the cell, the conversion of lactose into allolactose and then into glucose and lactose and finally the sequestration of repressor R by allolactose:

$$\begin{aligned}
\texttt{Transport:} \quad & \mathrm{L^e + P} \rightleftharpoons \mathrm{L^eP} \longrightarrow \mathrm{L + P} \\
\texttt{Conversion:} \quad & \mathrm{L + B} \rightleftharpoons \mathrm{LB} \longrightarrow \mathrm{A + B} \\
\texttt{Conversion:} \quad & \mathrm{A + B} \rightleftharpoons \mathrm{AB} \longrightarrow \mathrm{Glu + Gal + B} \\
\texttt{Sequestration:} \quad & \mathrm{A + R} \rightleftharpoons \mathrm{AR}
\end{aligned}$$

We see that the dynamics involve a number of enzymatic reactions and hence we can use Michaelis-Menten kinetics to model the response at a slightly reduced level of detail.

Given these reactions, we can write the reaction rate equations to describe the time evolution of the various species concentrations. Let α_{X} and K_{X} represent the parameters of the Michaelis-Menten functions and γ_{X} represent the dilution and

degradation rate for a given species X. The differential equation for the internal lactose concentration L becomes

$$\frac{dL}{dt} = \alpha_{\mathrm{LL^e}} P \frac{L^{\mathrm{e}}}{K_{\mathrm{L^e}} + L^{\mathrm{e}}} - \alpha_{\mathrm{AL}} B \frac{L}{K_{\mathrm{AL}} + L} - \gamma_L L, \tag{5.1}$$

where the first term arises from the transport of lactose into the cell, the second term is the conversion of lactose to allolactose, and the final term is due to degradation and dilution. Similarly, the dynamics for the allolactose concentration can be modeled as

$$\frac{dA}{dt} = \alpha_{\mathrm{AL}} B \frac{L}{K_{\mathrm{AL}} + L} - \alpha_{\mathrm{AB}} B \frac{A}{K_A + A} + k^{\mathrm{r}}_{\mathrm{AR}}[\mathrm{AR}] - k^{\mathrm{f}}_{\mathrm{AR}}[\mathrm{A}][\mathrm{R}] - \gamma_A A.$$

The dynamics of the production of β-gal and lactose permease are given by the transcription and translational dynamics of protein production. These genes are both part of the same operon (along with *lacA*) and hence the use a single mRNA strand for translation. To determine the production rate of mRNA, we need to determine the amount of repression that is present as a function of the amount of repressor, which in turn depends on the amount of allolactose that is present. We make the simplifying assumption that the sequestration reaction is fast, so that it is in equilibrium and hence

$$[\mathrm{AR}] = k_{\mathrm{AR}}[\mathrm{A}][\mathrm{R}], \qquad k_{\mathrm{AR}} = k^{\mathrm{f}}_{\mathrm{AR}} / k^{\mathrm{r}}_{\mathrm{AR}}.$$

We also assume that the total repressor concentration is constant (production matches degradation and dilution). Letting $R_{\mathrm{tot}} = [\mathrm{R}] + [\mathrm{AR}]$ represent the total repressor concentration, we can write

$$[\mathrm{R}] = R_{\mathrm{tot}} - k_{\mathrm{AR}}[\mathrm{A}][\mathrm{R}] \qquad \Longrightarrow \qquad [\mathrm{R}] = \frac{R_{\mathrm{tot}}}{1 + k_{\mathrm{AR}}[\mathrm{A}]}. \tag{5.2}$$

The simplification that the sequestration reaction is in equilibrium also simplifies the reaction dynamics for allolactose, which becomes

$$\frac{dA}{dt} = \alpha_{\mathrm{AL}} B \frac{L}{K_{\mathrm{AL}} + L} - \alpha_{\mathrm{A}} B \frac{A}{K_A + A} - \gamma_A A. \tag{5.3}$$

We next need to compute the effect of the repressor on the production of β-gal and lactose permease. It will be useful to express the promoter state in terms of the allolactose concentration A rather than R, using equation (5.2). We model this using a Hill function of the form

$$F_{\mathrm{BA}}(A) = \frac{\alpha_{\mathrm{R}}}{K_{\mathrm{R}} + R^n} = \frac{\alpha_{\mathrm{R}}(1 + K_{\mathrm{AR}} A)^n}{K_{\mathrm{R}}(1 + K_{\mathrm{AR}} A)^n + R^n_{\mathrm{tot}}}$$

Table 5.1: Parameter values for *lac* dynamics (from [99]).

Parameter	Value	Description
$\bar{\mu}$	3.03×10^{-2} min^{-1}	dilution rate
α_M	997 nMmin^{-1}	production rate of β-gal mRNA
β_B	1.66×10^{-2} min^{-1}	production rate of β-galactosidase
β_P	??? min^{-1}	production rate of lactose permease
α_A	1.76×10^{4} min^{-1}	production rate of allolactose
$\bar{\delta}_M$	0.411 min^{-1}	degradation and dilution of β-gal mRNA
$\bar{\gamma}_B$	8.33×10^{-4} min^{-1}	degradation and dilution of β-gal
$\bar{\gamma}_P$	?? min^{-1}	degradation and dilution of lactose permease
$\bar{\gamma}_A$	1.35×10^{-2} min^{-1}	degradation and dilution of allolactose
n	2	Hill coefficient for repressor
K	7200	
k_1	2.52×10^{-2} (μM)$^{-2}$	
K_L	0.97 μM	
K_A	1.95 μM	
β_A	2.15×10^{4} min^{-1}	
τ_M	0.10 min	
τ_B	2.00 min	
τ_P	??? min	

Letting M represent the concentration of the (common) mRNA, the resulting form of the protein production dynamics becomes

$$\begin{aligned}
\frac{dM}{dt} &= e^{-\mu\tau_M} F_{\mathrm{BA}}(A(t-\tau_m)) - \bar{\delta}_M M, \\
\frac{dB}{dt} &= \beta_B e^{-\mu\tau_B} M(t-\tau_B) - \bar{\gamma}_B B, \\
\frac{dP}{dt} &= \beta_P e^{-\mu(\tau_M+\tau_P)} M(t-\tau_M-\tau_P) - \bar{\gamma}_P P.
\end{aligned} \tag{5.4}$$

This model includes the degradation and dilution of mRNA ($\bar{\delta}_M$), the transcriptional delays β-gal mRNA (τ_M), the degradation and dilution of the proteins ($\bar{\gamma}_B$, $\bar{\gamma}_P$) and the delays in the translation and folding of the final proteins (τ_B, τ_P).

To study the dynamics of the circuit, we consider a slightly simplified situation in which we study the response to the internal lactose concentration L. In this case, we can take $L(t)$ as a constant and ignore the dynamics of the permease P. Figure 5.2a shows the time response of the system for an internal lactose concentration of 100 μM. As a test of the effect of time delays, we consider in Figure 5.2b the case when we set the delays τ_M and τ_B to both be zero. We see that the response has very little difference, consistent with our intuition that the delays are short compared to the dynamics of the underlying processes.

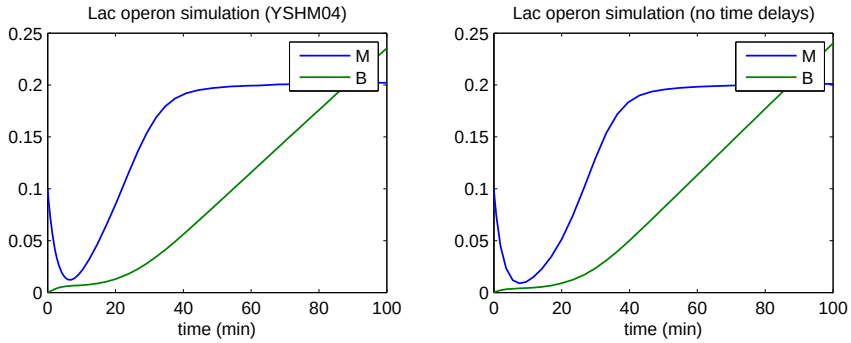


Figure 5.2: Time response of the Lac system.

Bifurcation analysis

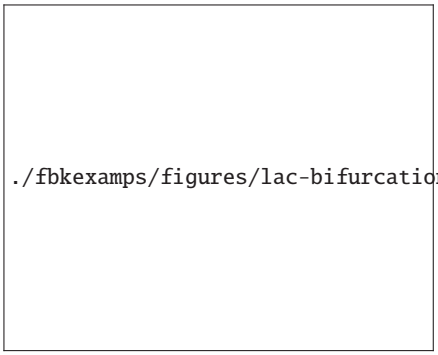
To further explore the different types of dynamics that can be exhibited by the Lac system, we make use of bifurcation analysis. If we vary the amount of lactose present in the environent, we expect that the lac circuitry will turn on at some point. Figure 5.3a shows the concentration of allolactose A as a function of the internal lactose concentration L. We see that the behavior of the *lac* system depends on the amount of lactose that is present in the cell. At low concentrations of lactose, the *lac* operon is turned off and the proteins required to metabolize lactose are not expressed. At high concentrations of lactose, the *lac* operon is turned on and the metabolic machinery is activated. In our model, these two operating conditions are measured by the concentration of β-galactosidase B and allolactose A. At intermediate concentrations of lactose, the system has multiple equilibrium points, with two stable equilibrium points corresponding to high and low concentrations of A (and B, as can be verified separately).

The parametric stability plot in Figure 5.3b shows the different types of behavior that can result based on the dilution rate μ and the lactose concentration L. We see that we get bistability only in a certain range of these parameters. Otherwise, we get that the circuitry is either uninduced or induced.

Sensitivity analysis

We now explore how the equilibrium conditions vary if the parameters in our model are changed.

For the gene *lacZ* (which encodes the protein β-galactosidase), we let B represent the protein concentration and M represent the mRNA concentration. We also consider the concentration of the lactose L inside the cell, which we will treat as an external input, and the concentration of allolactose, A. Assuming that the time delays considered previously can be ignored, the dynamics in terms of these variables



(a) Bifurcation diagram

(b) Stability diagram

Figure 5.3: Bifurcation and stability diagram for the lac system. Figures from [100].

are

$$\begin{aligned}
\frac{dM}{dt} &= F_{BA}(A,\theta) - \delta_b M, & F_{BA}(A,\theta) &= \alpha_{AB}\frac{1+k_1A^n}{K+k_1A^n},\\
\frac{dB}{dt} &= \beta_B M - \gamma_B B, & F_{AL}(L,\theta) &= \alpha_A\frac{L}{k_L+L}, \qquad (5.5)\\
\frac{dA}{ddt} &= BF_{AL}(L,\theta) - BF_{AA}(A,\theta) - \delta_A A, & F_{AA}(A,\theta) &= \beta_A\frac{A}{k_A+A}.
\end{aligned}$$

Here the state is $x = (M,B,A) \in \mathbb{R}^3$, the input is $w = L \in \mathbb{R}$ and the parameters are $\theta = (\alpha_B,\beta_B,\alpha_A,\delta_B,\gamma_B,\delta_A,n,k,k_1,k_L,k_A,\beta_A) \in \mathbb{R}^{12}$. The values for the parameters are listed in Table 5.1.

We investigate the dynamics around one of the equilibrium points, corresponding to an intermediate input of $L = 40\ \mu$M. There are three equilibrium points at this value of the input:

$$x_{1,e} = (0.000393, 0.000210, 3.17), \qquad x_{2,e} = (0.00328, 0.00174, 19.4), \qquad x_{3,e} = (0.0142, 0.00758, 42.1).$$

We choose the third equilibrium point, corresponding to the lactose metabolic machinery being activitated and study the sensitivity of the steady state concentrations of allolactose (A) and β-galactosidase (B) to changes in the parameter values.

The dynamics of the system can be represented in the form $dx/dt = f(x,\theta,L)$ with

$$f(x,\theta,L) = \begin{pmatrix} F_{\mathrm{BA}}(A,\theta) - \delta_B M - \mu M \\ \beta_B M - \gamma_B B - \mu B \\ F_{\mathrm{AL}}(L,\theta)B - F_{\mathrm{AA}}(A,\theta)B - \gamma_A A - \mu A \end{pmatrix}.$$

To compute the sensitivity with respect to the parameters, we compute the deriva-

tives of f with respect to the state x,

$$\frac{\partial f}{\partial x} = \begin{pmatrix} -\delta_B - \mu & 0 & \frac{\partial F_{BA}}{\partial A} \\ \beta_B & -\gamma_B - \mu & 0 \\ 0 & F_{AL} - F_{\mathrm{AA}} & -B\frac{\partial F_{AA}}{\partial A} \end{pmatrix}$$

and the parameters θ,

$$\frac{\partial f}{\partial \theta} = \begin{pmatrix} F_{BA} & 0 & 0 & -M & 0 & 0 & \frac{\partial F_{BA}}{\partial n} & \frac{\partial F_{BA}}{\partial k} & \frac{\partial F_{BA}}{\partial k_1} & 0 & 0 & 0 \end{pmatrix}.$$

Carrying out the relevant computations and evaluating the resulting expression numerically, we obtain

$$\frac{\partial}{\partial \theta}\begin{pmatrix} B_e \\ A_e \end{pmatrix} = \begin{pmatrix} -1.21 & 0.0243 & -3.35\times 10^{-6} & 0.935 & 1.46 & \dots & 0.00115 \\ -2720. & 47.7 & -0.00656 & 1830. & 2860. & \dots & 3.27 \end{pmatrix}.$$

We can also normalize the sensitivity computation, as described in equation (3.9):

$$\bar{S}_{x_e\theta} = \frac{\partial x_e / x_e}{\partial \theta / \theta_0} = (D^x)^{-1} S_{x_e\theta} D^\theta,$$

where $D^x = \mathrm{diag}\{x_e\}$ and $D^\theta = \mathrm{diag}\{\theta_0\}$, which yields

$$\bar{S}_{y_e\theta} = \begin{pmatrix} -4.85 & 3.2 & -3.18 & 3.11 & 3.2 & 6.3 & -6.05 & -4.1 & 4.02 & 6.05 \\ -1.96 & 1.13 & -1.12 & 1.1 & 1.13 & 3.24 & -3.11 & -2.11 & 2.07 & 3.11 \end{pmatrix}$$

where

$$\theta = \begin{pmatrix} \mu & \alpha_M & K & K_1 & \beta_B & \alpha_A & K_L & \beta_A & K_A & L \end{pmatrix}.$$

We see from this computation that increasing the growth rate decreases the equilibrium concentation of B and A, while increasing the lactose concentration by 2-fold increases the equilibrium β-gal concentration 12-fold (6X) and the allolactose concentration by 6-fold (3X).

5.2 Bacterial Chemotaxis

Chemotaxis refers to the process by which micro-organisms move in response to chemical stimuli. Examples of chemotaxis include the ability of organisms to move in the direction of nutrients or move away from toxins in the environment. Chemotaxis is called *positive chemotaxis* if the motion is in the direction of the stimulus and *negative chemotaxis* if the motion is away from the stimulant, as shown in Figure 5.4. Many chemotaxis mechanisms are stochastic in nature, with biased random motions causing the average behavior to be either positive, negative or neutral (in the absence of stimuli).

In this section we look in some detail at bacterial chemotaxis, which *E. coli* use to move in the direction of increasing nutrients. The material in this section is based primarily on the work of Barkai and Leibler [8] and Rao, Kirby and Arkin [77].

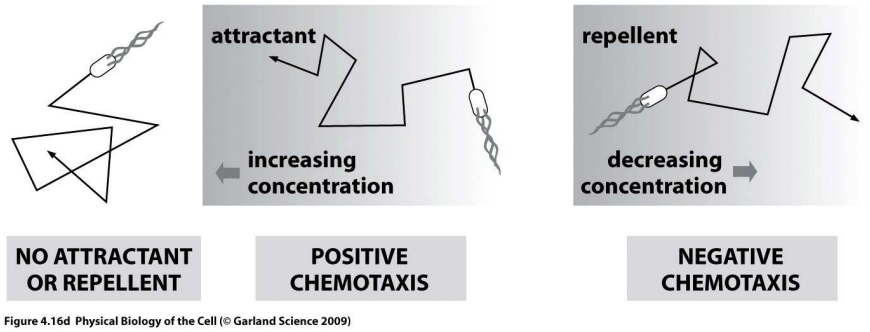


Figure 5.4: Examples of chemotaxis. Figure from Phillips, Kondev and Theriot [72]; used with permission of Garland Science.

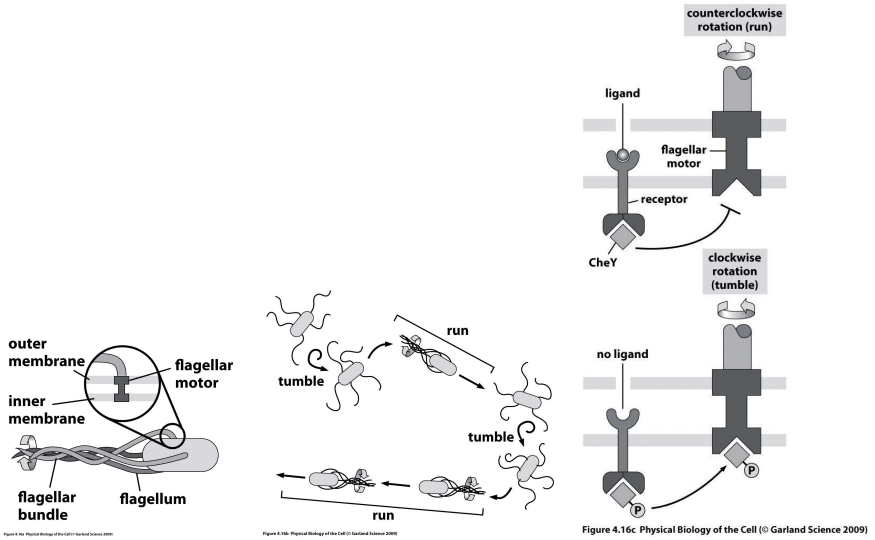
Control system overview

The chemotaxis system in *E. coli* consists of a sensing system that detects the presence of nutrients, and actuation system that propels the organism in its environment, and control circuitry that determines how the cell should move in the presence of chemicals that stimulate the sensing system.

The actuation system in the *E. coli* consists of a set of flagella that can be spun using a flagellar motor embedded in the outer membrane of the cell, as shown in Figure 5.5a. When the flagella all spin in the counter clockwise direction, the individual flagella form a bundle and cause the organism to move roughly in a straight line. This behavior is called a "run" motion. Alternatively, if the flagella spin in the clockwise direction, the individual flagella do not form a bundle and the organism "tumbles", causing it to rotate (Figure 5.5b). The selection of the motor direction is controlled by the protein CheY: if phosphorylated CheY binds to the motor complex, the motor spins clockwise (tumble), otherwise it spins counter-clockwise (run).

Because of the size of the organism, it is not possible for a bacterium to sense gradients across its length. Hence, a more sophisticated strategy is used, in which the organism undergoes a combination of run and tumble motions. The basic idea is illustrated in Figure 5.5c: when high concentration of ligand (nutrient) is present, the CheY protein is left unphosphorylated and does not bind to the actuation complex, resulting in a counter-clockwise rotation of the flagellar motor (run). Conversely, if the ligand is not present then the molecular machinery of the cell causes CheY to be phosphorylated and this modifies the flagellar motor dynamics so that a clockwise rotation occurs (tumble). The net effect of this combination of behaviors is that when the organism is traveling through regions of higher nutrient concentration, it continues to move in a straight line for a longer period before tumbling, causing it to move in directions of increasing nutrient concentration.

A simple model for the molecular control system that regulates chemotaxis is



(a) (b) (c)

Figure 5.5: Bacterial chemotaxis. Figures from Phillips, Kondev and Theriot [72]; used with permission of Garland Science.

shown in Figure 5.6. We start with the basic sensing and and actuation mechanisms. A membrane bound protein MCP (methyl-accepting chemotaxis protein) that is capable of binding to the external ligand serves as a signal transducing element from the cell exterior to the cytoplasm. Two other proteins, CheW and CheA, form a complex with MCP. This complex can either be in an active or inactive state. In the active state, CheA is autophosphorylated and serves as a phosphotransferase for two additional proteins, CheB and CheY. The phosphorylated form of CheY then binds to the motor complex, causing clockwise rotation of the motor.

The activity of the receptor complex is governed by two primary factors: the binding of a ligand molecule to the MCP protein and the presence or absence of up to 4 methyl groups on the MCP protein. The specific dependence on each of these factors is somewhat complicated. Roughly speaking, when the ligand L is bound to the receptor then the complex is less likely to be active. Furthermore, as more methyl groups are present, the ligand binding probability increases, allowing the gain of the sensor to be adjusted through methylation. Finally, even in the absence of ligand the receptor complex can be active, with the probability of it being active increasing with increased methylation. Figure 5.7 summarizes the possible states, their free energies and the probability of activity.

Several other elements are contained in the chemotaxis control circuit. The most important of these are implemented by the proteins CheR and CheB, both of which affect the receptor complex. CheR, which is constitutively produced in the cell, methylates the receptor complex at one of the four different methylation sites. Con-

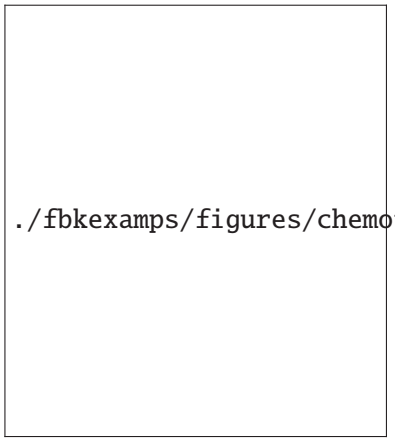


Figure 5.6: Control system for chemotaxis. Figure from Rao *et al.* [77] (Figure 1A).

versely, the phosphorylated form of CheB demethylates the receptor complex. As described above, the methylation patterns of the receptor complex affect its activity, which affects the phosphorylation of CheA and, in turn, phosphorylation of CheY and CheB. The combination of CheA, CheB and the methylation of the receptor complex forms a negative feedback loop: if the receptor is active, then CheA phosphorylates CheB, which in turn demethylates the receptor complex, making it less active. As we shall see when we investigate the detailed dynamics below, this feedback loop corresponds to a type of integral feedback law. This integral action allows the cell to adjust to different levels of ligand concentration, so that the behavior of the system is invariant to the absolute nutrient levels.

Modeling

The detailed reactions that implement chemotaxis are illustrated in Figure 5.8. Letting T represent the receptor complex and T^A represent an active form, the basic reactions can be written as

$$\begin{aligned} T^A + A &\rightleftharpoons T^A{:}A \longrightarrow A^p + T^A \\ A^p + B &\rightleftharpoons A^p{:}B \longrightarrow A + B^p \qquad B^p + P \rightleftharpoons B^p{:}P \longrightarrow B + P \\ A^p + Y &\rightleftharpoons A^p{:}Y \longrightarrow A + Y^p \qquad Y^p + Z \rightleftharpoons Y^p{:}Z \longrightarrow Y + Z \end{aligned} \tag{5.6}$$

where CheA, CheB, CheY and CheZ are written simply as A, B, Y and Z for simplicity and *P* is a non-specific phosphatase. We see that these are basically three linked sets of phosphorylation and dephosphorylation reactions, with CheA serving as a phosphotransferase and P and CheZ serving as phosphatases.

The description of the methylation of the receptor complex is a bit more complicated. Each receptor complex can have multiple methyl groups attached and the

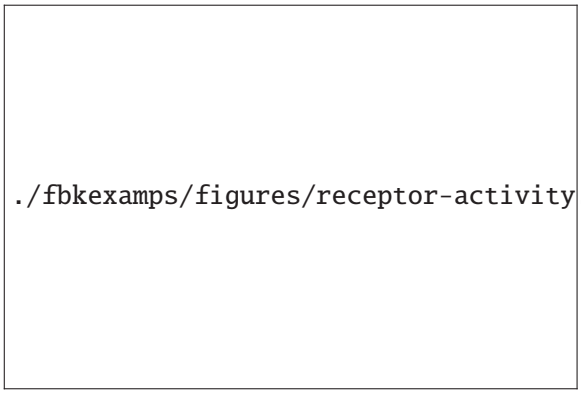


Figure 5.7: Receptor complex states. The probability of a given state being in an active configuration is given by p. Figure obtained from [66].

activity of the receptor complex depends on both the amount of methylation and whether a ligand is attached to the receptor site. Furthermore, the binding probabilities for the receptor also depend on the methylation pattern. To capture this, we use the set of reactions that are illustrated in Figures 5.6 and 5.8. In this diagram, $\mathrm{T}_\mathrm{i}^\mathrm{s}$ represents a receptor that has i methylation sites filled and ligand state s (which can be either u if unoccupied or o if occupied). We let M represent the maximum number of methylation sites ($M = 4$ for *E. coli*).

Using this notation, the transitions between the states correspond to the reactions shown in Figure 5.9:

$$\begin{aligned}
&\mathrm{T}_\mathrm{i}^\mathrm{x} + \mathrm{B}^\mathrm{p} \rightleftharpoons \mathrm{T}_\mathrm{i}^\mathrm{x}{:}\mathrm{B}^\mathrm{p} \longrightarrow \mathrm{T}_{\mathrm{i}-1}^\mathrm{x} + \mathrm{B}^\mathrm{p} && \mathrm{i} > 0,\ x \in \{u, 0\} \\
&\mathrm{T}_\mathrm{i}^\mathrm{x} + \mathrm{R} \rightleftharpoons \mathrm{T}_\mathrm{i}^\mathrm{x}{:}\mathrm{R} \longrightarrow \mathrm{T}_{\mathrm{i}+1}^\mathrm{x} + \mathrm{R} && \mathrm{i} < M,\ x \in \{u, 0\} \\
&\mathrm{T}_\mathrm{i}^\mathrm{u} + \mathrm{L} \rightleftharpoons \mathrm{T}_\mathrm{i}^\mathrm{o}
\end{aligned}$$

We now must write reactions for each of the receptor complexes with CheA. Each form of the receptor complex has a different activity level and so the most complete description is to write a separate reaction for each $\mathrm{T}_\mathrm{i}^\mathrm{o}$ and $\mathrm{T}_\mathrm{i}^\mathrm{u}$ species:

$$\mathrm{T}_\mathrm{i}^\mathrm{x} + \mathrm{A} \underset{k_i^{r,o}}{\overset{k_i^{f,o}}{\rightleftharpoons}} \mathrm{T}_\mathrm{i}^\mathrm{x}{:}\mathrm{A} \xrightarrow{k_i^{c,o}} \mathrm{A}^\mathrm{p} + \mathrm{T}_\mathrm{i}^\mathrm{x},$$

where $\mathrm{x} \in \{\mathrm{o}, \mathrm{u}\}$ and $\mathrm{i} = 0, \ldots, M$. This set of reactions replaces the placeholder reaction $\mathrm{T}^\mathrm{A} + \mathrm{A} \rightleftharpoons \mathrm{T}^\mathrm{A}{:}\mathrm{A} \longrightarrow \mathrm{A}^\mathrm{p} + \mathrm{T}^\mathrm{A}$ used earlier.

Approximate model

The detailed model described above is sufficiently complicated that it can be difficult to analyze. In this section we develop a slightly simpler model that can be

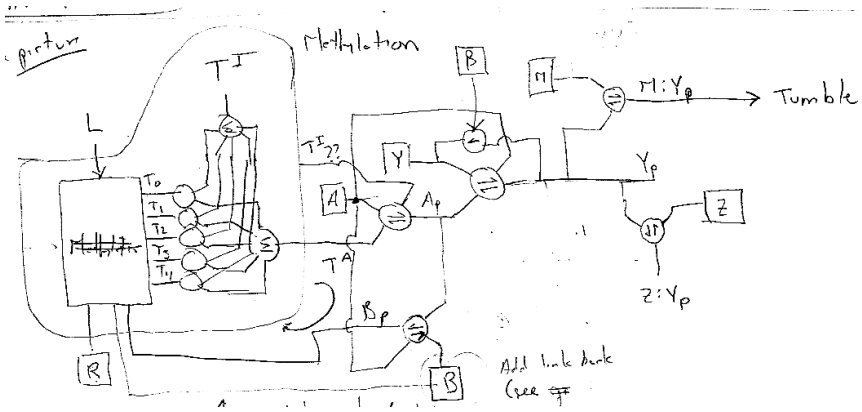


Figure 5.8: Circuit diagram for chemotaxis.

used to explore the adaptation properties of the circuit, which happen on a slower time-scale.

We begin by simplifying the representation of the receptor complex and its methylation pattern. Let $L(t)$ represent the ligand concentration and T_i represent the concentration of the receptor complex with i sides methylated. If we assume that the binding reaction of the ligand L to the complex is fast, we can write the probability that a receptor complex with i sites methylated is in its active state as a static function $\alpha_i(L)$, which we take to be of the form

$$\alpha_i(L) = \frac{\alpha_i^o L}{K_L + L} + \frac{\alpha_i K_L}{K_L + L}.$$

The coefficients α_i^o and α_i capture the effect of presence or absence of the ligand on the activity level of the complex. Note that α_i has the form of a Michaelis-Menten function, reflecting our assumption that ligand binding is fast compared to the rest

./fbkexamps/figures/chemotaxis-methylation.eps

Figure 5.9: Methylation model for chemotaxis. Figure from Barkai and Leibler [8] (Box 1). Note: the figure uses the notation E_i^s for the receptor complex instead of T_i^s.

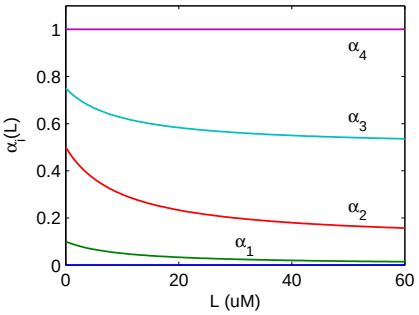


Figure 5.10: Probability of activity.

of the dynamics in the model. Following [77], we take the coefficients to be

$$a_0 = 0, \quad a_1 = 0.1, \quad a_2 = 0.5, \quad a_3 = 0.75, \quad a_4 = 1,$$
$$a_0^o = 0, \quad a_1^o = 0, \quad a_2^o = 0.1, \quad a_3^o = 0.5, \quad a_4^o = 1.$$

and choose $K_L = 10$ μM. Figure 5.10 shows how each α_i varies with L.

The total concentration of active receptors can now be written in terms of the receptor complex concentrations T_i and the activity probabilities $\alpha_i(L)$. We write the concentration of activated complex T^{A} and inactivated complex T^{I} as

$$T^A = \sum_{i=0}^{4} \alpha_i(L) T_i, \qquad T^I = \sum_{i=0}^{4} (1 - \alpha_i(L)) T_i.$$

These formulas can now be used in our dynamics as an effective concentration of active or inactive receptors, justifying the notation that we used in equation (5.6).

We next model the transition between the methylation patterns on the receptor. We assume that the rate of methylation depends on the activity of the receptor complex, with active receptors less likely to be demethylated and inactive receptors less likely to be methylated [77, 66]. Let

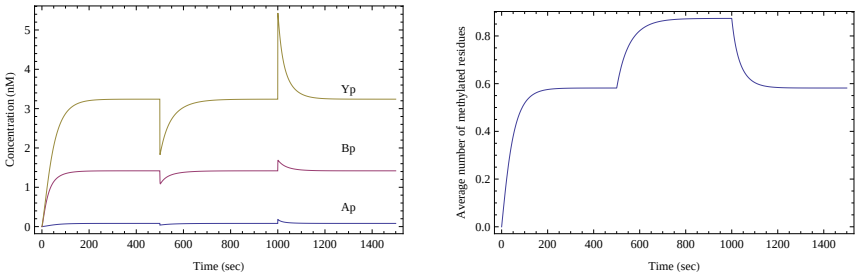
$$r_B = k_B \frac{B^p}{K_B + T^A}, \qquad r_R = k_R \frac{R}{K_R + T^I},$$

represent rates of the methylation and demethylation reactions. We choose the coefficients as

$$k_B = 0.5, \quad K_B = 5.5, \quad k_R = 0.255, \quad K_R = 0.251,$$

We can now write the methylation dynamics as

$$\frac{d}{dt} T_i = r_R(1 - \alpha_{i+1}(L)) T_{i-1} + r_B \alpha_{i+1}(L) T_{i+1} - r_R(1 - \alpha_i(L)) T_i - r_B \alpha_i(L) T_i,$$



(a) (b)

Figure 5.11: Simulation and analysis of reduced-order chemotaxis model.

where the first and second terms represent transitions into this state via methylation or demethylation of neighboring states (see Figure 5.9) and the last two terms represent transitions out of the current state by methylation and demethylation, respectively. Note that the equations for T_0 and T_4 are slightly different since the demethylation and methylation reactions are not present, respectively.

Finally, we write the dynamics of the phosphorylation and dephosphorylation reactions, and the binding of $\mathrm{CheY^p}$ to the motor complex. Under the assumption that the concentrations of the phosphorylated proteins are small relative to the total protein concentrations, we can approximate the reaction dynamics as

$$\begin{aligned}
\frac{d}{dt}A^p &= 50T^A A - 100A^pY - 30A^pB, \\
\frac{d}{dt}Y^p &= 100A^pY - 0.1Y^p - 5[\mathrm{M}]Y^p + 19[\mathrm{M{:}Y^p}] - 30Y^p, \\
\frac{d}{dt}B^p &= 30A^pB - B^p, \\
\frac{d}{dt}[\mathrm{M{:}Y^p}] &= 5[\mathrm{M}]Y^p - 19[\mathrm{M{:}Y^p}].
\end{aligned}$$

The total concentrations of the species are given by

$$A + A^p = 5\ \mathrm{nM}, \qquad B + B^p = 2\ \mathrm{nM}, \qquad Y + Y^p + [\mathrm{M{:}Y^p}] = 17.9\ \mathrm{nM},$$
$$[\mathrm{M}] + [\mathrm{M{:}Y^p}] = 5.8\ \mathrm{nM}, \qquad R = 0.2\ \mathrm{nM}, \qquad \textstyle\sum_{i=0}^{4} T_i = 5\ \mathrm{nM}.$$

The reaction coefficients and concentrations are taken from Rao *et al.* [77].

Figure 5.11a shows a the concentration of the phosphorylated proteins based on a simulation of the model. Initially, all species are started in their unphosphorylated and demethylated states. At time $T = 500$ s the ligand concentration is increased to $L = 10$ μM and at time $T = 1000$ it is returned to zero. We see that immediately after the ligand is added, the $\mathrm{CheY^p}$ concentration drops, allowing longer runs between tumble motions. After a short period, however, the $\mathrm{CheY^p}$ concentration adapts to the higher concentration and the nominal run versus tumble behavior is restored.

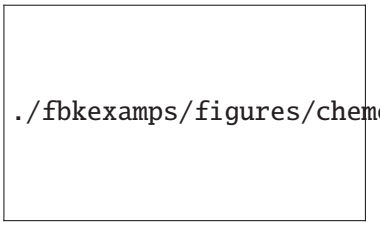


Figure 5.12: Reduced order model of receptor activity. Obtained from [3], Figure 7.9.

Similarly, after the ligand concentration is decreased the concentration of $\mathrm{CheY}^{\mathrm{p}}$ increases, causing a larger fraction of tumbles (and subsequent changes in direction). Again, adaptation over a longer time scale returns that CheY concentration to its nominal value.

Figure 5.11b helps explain the adaptation response. We see that the average amount of methylation of the receptor proteins increases when the ligand concentration is high, which decreases the activity of CheA (and hence decreases the phosphorylation of CheY).

Integral action

The perfect adaptation mechanism in the chemotaxis control circuitry has the same function as the use of integral action in control system design: by including a feedback on the integral of the error, it is possible to provide exact cancellation to constant disturbances. In this section we demonstrate that a simplified version of the dynamics can indeed be regarded as integral action of an appropriate signal. This interpretation was first pointed out by Yi *et al* [98].

We begin by formulating an even simpler model for the system dynamics that captures the basic features required to understand the integral action. Let X represent the receptor complex and assume that it is either methylated or not. We let $\mathrm{X_m}$ represent the methylated state and we further assume that this methylated state can be activated, which we write as $\mathrm{X_m^*}$. This simplified description replaces the multiple states T_i and probabilities $\alpha_i(L)$. We also ignore the additional phosphorylation dynamics of CheY and simply take the activated receptor concentration X_m^* as our measure of overall activity.

Figure 5.12 shows the transitions between the various forms X. As before, CheR methylates the receptor and $\mathrm{CheB}^{\mathrm{p}}$ demethylates it. We simplify the picture by only allowing $\mathrm{CheB}^{\mathrm{p}}$ to act on the active state $\mathrm{X_m^*}$ and CheR to act on the inactive state. We take the ligand into account by assuming that the transition between the active form $\mathrm{X_m^*}$ and the inactive form $\mathrm{X_m}$ depends on the ligand concentration: higher ligand concentration will increase the rate of transition to the inactive state.

This model is a considerable simplification from the ligand binding model that is illustrated in Figures 5.7 and 5.9. In the previous models, there is some prob-

ability of activity with or without methylation and with or without ligand. In this simplified model, we assume that only three states are of interest: demethylated, methylated/inactive and methylated/active. We also modify the way that that ligand binding is captured and instead of keeping track of all of the possibilities in Figure 5.7, we assume that the ligand transitions us from an active state $\mathrm{X_m^*}$ to an inactive $\mathrm{X_m}$. These states and transitions are roughly consistent with the different energy levels and probabilities in Figure 5.7, but it is clearly a much coarser model.

Accepting these approximations, the model illustrated in Figure 5.12 results in a set of chemical reactions of the form

$$\begin{aligned}
&\texttt{R1}: && \mathrm{X + R \rightleftharpoons X{:}R \longrightarrow X_m + R} && \text{methylation}\\
&\texttt{R2}: && \mathrm{X_m^* + B^p \rightleftharpoons X_m^*{:}B^p \longrightarrow X + B^p} && \text{demethylation}\\
&\texttt{R3}: && \mathrm{X_m^* \underset{k^r}{\overset{k^f(L)}{\rightleftharpoons}} X_m} && \text{activation/deactivation}
\end{aligned}$$

For simplicity we take both R and $\mathrm{B^p}$ to have constant concentration.

We can approximate the first and second reactions by their Michaelis-Menten forms, which yield net methylation and demethylation rates (for those reactions)

$$v_+ = k_R R \frac{X}{K_X + X}, \qquad v_- = k_B B^p \frac{X_m^*}{K_{X_m^*} + X_m^*}.$$

If we further assume that $X \gg K_X > 1$, then the methylation rate can be further simplified:

$$v_+ = k_R R \frac{X}{K_X + X} \approx K_R R.$$

Using these approximations, we can write the resulting dynamics for the overall system as

$$\begin{aligned}
\frac{d}{dt} X_m &= k_R R + k^f(L) X_m^* - k^r X_m \\
\frac{d}{dt} X_m^* &= -k_B B^p \frac{X_m^*}{K_{X_m^*} + X_m^*} - k^f(L) X_m^* + k^r X_m.
\end{aligned}$$

We wish to use this model to understand how the steady state activity level X_m^* depends on the ligand concentration L (which enters through the deactivation rate $k^f(L)$).

It will be useful to rewrite the dynamics in terms of the activated complex concentration X_m^* and the *total* methylated complex concentration $X_m^t = X_m + X_m^*$. A simple set of algebraic manipulations yields

$$\begin{aligned}
\frac{dX_m^*}{dt} &= k^r (X_m^t - X_m^*) - k_B B^p \frac{X_m^*}{K_{X_m^*} + X_m^*} - k^f(L) X_m^*, \\
\frac{dX_m^t}{dt} &= k_R R - k_B B^p \frac{X_m^*}{K_{X_m^*} + X_m^*}.
\end{aligned}$$

From the second equation, we see that the the concentration of methylated complex $\mathrm{X_m^t}$ is a balance between the action of the methylation reaction (R1, characterized by v_+) and the demethylation reaction (R2, at rate v_-). Since the action of a ligand binding to the receptor complex increases the rate of deactivation of the complex (R3), in the presence of a ligand we will increase the amount of methylated complex (and, via reaction R1) eventually restore the amount of the activated complex. This represents the adaptation mechanism in this simplified model.

To further explore the effect of adaptation, we compute the equilibrium points for the system. Setting the time derivatives to zero, we obtain

$$X^*_{m,e} = \frac{K_{X^*_m} k_R R}{k_B B^p - k_R R}$$
$$X^t_{m,e} = \frac{1}{k^r}\left(k^r X^*_m + k_B B^p \frac{X^*_m}{K_{X^*_m} + X^*_m} + k^f(L) X^*_m\right).$$

Note that the solution for the active complex $X^*_{m,e}$ in the first equation does not depend on $k^f(L)$ (or k^r) and hence the steady state solution is independent of the ligand concentration. Thus, in steady state, the concentration of activated complex adapts to the steady state value of the ligand that is present, making it insensitive to the steady state value of this input.

The dynamics for X^t_m can be viewed as an integral action: when the concentration of X^*_m matches its reference value (with no ligand present), the quantity of methylated complex X^t_m remains constant. But if X^t_m does not match this reference value, then X^t_m increases at a rate proportional to the methylation "error" (measured here by difference in the nominal reaction rates v_+ and v_-). It can be shown that this type of integral action is necessary to achieve perfect adaptation in a robust manner [98].

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.

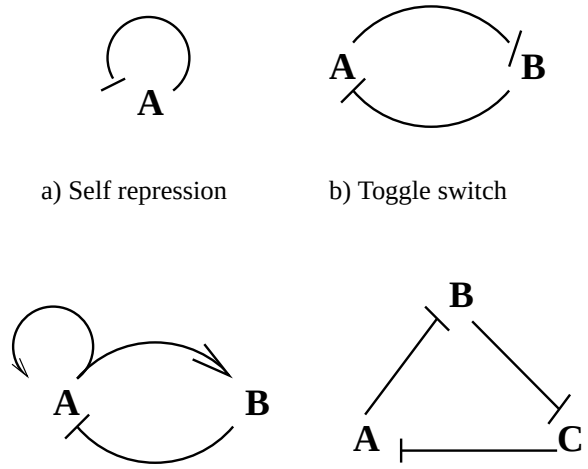
6.1 Introduction to Biological Circuit Design

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 [23], a



a) Self repression b) Toggle switch

c) Activator-repressor clock d) Repressilator

Figure 6.1: Early gene circuits that have been fabricated in bacteria *E. coli*: the negatively autoregulated gene [9], the toggle switch [27], the activator-repressor clock [5], and the repressilator [23].

toggle switch obtained by connecting two inverters in a ring fashion [27], and an autorepressed circuit [9] (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 to 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 in 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 negatively autoregulated system is given by

$$\frac{dA}{dt} = \frac{\beta}{1+(A/K)^n} - \gamma A. \tag{6.1}$$

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

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

in which β_0 is the unrepressed production rate. We refer to this system as the open loop system.

Dynamic effects of negative autoregulation

As we showed via simulation in Example 2.2, 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 $A_e = \beta_0/\gamma$ 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

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

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

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

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 $t_{\text{half}} = \ln(2)/\gamma$. Note that the half-life does not depend on the production rate β_0 and only depends on the protein decay rate constant γ.

Let now A_e be the steady state of the negatively autoregulated system (6.1). 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

$$\frac{dz}{dt} = -\overline{\gamma} z,$$

in which

$$\overline{\gamma} = \gamma + \beta \frac{n A_e^{n-1}/K^n}{(1+(A_e/K)^n)^2}.$$

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

Since $\overline{\gamma} > \gamma$ (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 2.2.

Noise filtering

In this section, we investigate the effect of the negative autoregulation on the noise of the system. In order to do this, we employ the Langevin modeling framework and determine the frequency response to the intrinsic noise on the various reactions. 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 can be well approximated by 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. This is the case, for example, when degradation tags are added to the protein to make its decay rate larger. 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 to noise.

Assuming that mRNA is at its quasi-steady state

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

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

in which β_0 is the constitutive production rate, p is the DNA promoter, and γ 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

$$\frac{dn_A}{dt} = \beta_0 n_p - \gamma n_A + \sqrt{\beta_0 n_p}\Gamma_1 - \sqrt{\gamma n_A}\Gamma_2,$$

in which Γ_1 and Γ_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

$$\frac{dA}{dt} = \beta_0 p_{\text{tot}} - \gamma A + \frac{1}{\sqrt{\Omega}}(\sqrt{\beta_0 p_{\text{tot}}}\Gamma_1 - \sqrt{\gamma A}\Gamma_2).$$

This is a linear system and therefore we can calculate the frequency response to any of the two inputs Γ_1 and Γ_2. The frequency response to input Γ_1 has magnitude given by

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

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

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

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

Employing the Langevin equation (4.9) of Section 4.1, we obtain

$$\begin{aligned}
\frac{dp}{dt} &= -aAp + d(p_{\text{tot}} - p) + \frac{1}{\sqrt{\Omega}}(-\sqrt{aAp}\Gamma_3 + \sqrt{d(p_{\text{tot}} - p)}\Gamma_4) \\
\frac{dA}{dt} &= \beta p - \gamma A - aAp + d(p_{\text{tot}} - p) + \frac{1}{\sqrt{\Omega}}(\sqrt{\beta p}\Gamma_1 - \sqrt{\gamma A}\Gamma_2 - \sqrt{aAp}\Gamma_3 + \\
&\quad \sqrt{d(p_{\text{tot}} - p)}\Gamma_4),
\end{aligned}$$

in which Γ_3 and Γ_4 are the noises on the association and on the dissociation reactions, respectively. Letting $K_{\text{d}} = d/a$, $N_1 = \frac{1}{\sqrt{\Omega}}(-\sqrt{Ap/K_{\text{d}}}\Gamma_3 + \sqrt{(p_{\text{tot}} - p)}\Gamma_4)$, and $N_2 = \frac{1}{\sqrt{\Omega}}(\sqrt{\beta p}\Gamma_1 - \sqrt{\gamma A}\Gamma_2)$, we can rewrite the above system in the following form:

$$\begin{aligned}
\frac{dp}{dt} &= -aAp + d(p_{\text{tot}} - p) + \sqrt{d}N_1(t) \\
\frac{dA}{dt} &= \beta p - \gamma A - aAp + d(p_{\text{tot}} - p) + N_2(t) + \sqrt{d}N_1(t).
\end{aligned}$$

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

$$\begin{aligned}
\epsilon\frac{dp}{dt} &= -\gamma Ap/K_{\text{d}} + \gamma(p_{\text{tot}} - p) + \sqrt{\epsilon}\sqrt{\gamma}N_1(t) \\
\frac{dy}{dt} &= \beta p - \gamma(y + p) + N_2(t).
\end{aligned}$$

By setting $\epsilon = 0$, we obtain the quasi-steady state value $p = p_{\text{tot}}/(A/K_{\text{d}} + 1)$. Writing $\dot{A} = \dot{y} + \dot{p}$, using the chain rule for $\dot{p}$, and assuming that $p_{\text{tot}}/K_{\text{d}}$ is sufficiently small, we obtain the reduced system describing the dynamics of A as

$$\frac{dA}{dt} = \beta \frac{p_{\text{tot}}}{A/K_{\text{d}} + 1} - \gamma A + \frac{1}{\sqrt{\Omega}}(\sqrt{\beta p}\Gamma_1 - \sqrt{\gamma A}\Gamma_2) =: f(A, \Gamma_1, \Gamma_2).$$

The equilibrium point for this system corresponding to the mean values $\Gamma_1 = 0$ and $\Gamma_2 = 0$ of the inputs is given by

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

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

$$\left.\frac{\partial f}{\partial A}\right|_{A_e, \Gamma_1 = 0, \Gamma_2 = 0} = -\beta \frac{p_{\text{tot}}/K_{\text{d}}}{(A_e/K_{\text{d}} + 1)^2} - \gamma =: -\overline{\gamma},$$

$$b_1 = \left.\frac{\partial f}{\partial \Gamma_1}\right|_{A_e, \Gamma_1 = 0, \Gamma_2 = 0} = \frac{1}{\sqrt{\Omega}} \sqrt{\frac{\beta p_{\text{tot}}}{A_e/K_{\text{d}} + 1}}, \qquad b_2 = \left.\frac{\partial f}{\partial \Gamma_2}\right|_{A_e, \Gamma_1 = 0, \Gamma_2 = 0} = -\frac{1}{\sqrt{\Omega}} \sqrt{\gamma A_e}.$$

Hence, the frequency response to Γ_1 has magnitude given by

$$M^c(\omega) = \frac{b_1}{\sqrt{\omega^2 + \overline{\gamma}^2}}.$$

In order to make a fair comparison between this response and that of the open loop system, we need the steady states of both systems to be the same. In order to guarantee this, 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, we have that $b_1 = \sqrt{\beta_0 p_{\text{tot}}/\Omega}$. Since we also have that $\overline{\gamma} > \gamma$, it follows that $M^c(\omega) < M(\omega)$ for all ω. That is, the gain of the closed loops system is smaller than that of the open loop system. This result implies that negative autoregulation attenuates noise at all frequencies. The two frequency responses are plotted in Figure 6.2(a).

mRNA decay close to protein decay

In the case in which mRNA and protein decay rates are comparable, we need to model both the processes of transcription and translation. Letting m_A denote the mRNA of A, the reactions describing the open loop system modify to

$$\text{R}_1\colon \text{m}_\text{A} \xrightarrow{\kappa} \text{m}_\text{A} + \text{A}, \qquad \text{R}_2\colon \text{A} \xrightarrow{\gamma} \emptyset, \qquad \text{R}_5\colon \text{p} \xrightarrow{\alpha_0} \text{m}_\text{A} + \text{p}, \qquad \text{R}_6\colon \text{m}_\text{A} \xrightarrow{\delta} \emptyset,$$

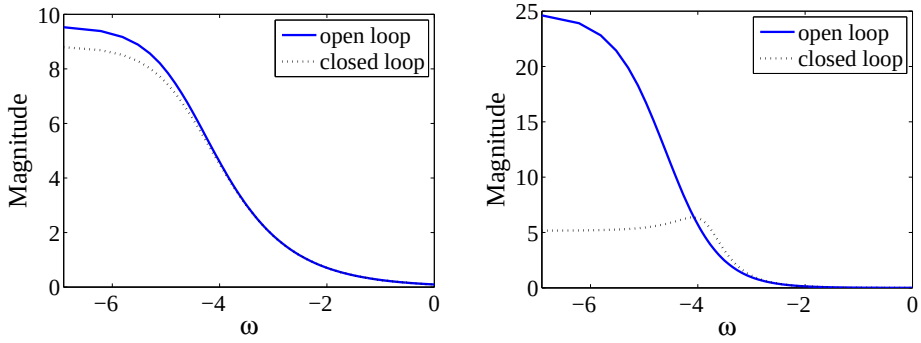


Figure 6.2: (a) Magnitude of the frequency response to noise $\Gamma_1(t)$ for both open loop and closed loop for the model in which mRNA is assumed at its quasi-steady state. The parameters are $p_{\text{tot}} = 10$, $K_{\text{d}} = 10$, $\beta = 0.001$, $\gamma = 0.01$, $\Omega = 1$, and $\beta_0 = 0.00092$. (b) Frequency response to noise $\Gamma_6(t)$ for both open loop and closed loop for the model in which mRNA decay is close to protein decay. The parameters are $p_{\text{tot}} = 10$, $K_{\text{d}} = 10$, $\alpha = 0.001$, $\beta = 0.01$, $\delta = 0.01$, $\gamma = 0.01$, and $\alpha_0 = 0.0618$.

while those describing the closed loop system modify to

$$\mathrm{R}_1\colon \mathrm{m_A} \xrightarrow{\kappa} \mathrm{m_A} + \mathrm{A}, \qquad \mathrm{R}_2\colon \mathrm{A} \xrightarrow{\gamma} \emptyset,$$

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

$$\mathrm{R}_5\colon \mathrm{p} \xrightarrow{\alpha} \mathrm{m_A} + \mathrm{p}, \qquad \mathrm{R}_6\colon \mathrm{m_A} \xrightarrow{\delta} \emptyset, \qquad p_{\text{tot}} = p + C.$$

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

$$\frac{dm_A}{dt} = F(A) - \delta m_A + \frac{1}{\sqrt{\Omega}}\left(\sqrt{F(A)}\Gamma_5 - \sqrt{\delta m_A}\Gamma_6\right)$$
$$\frac{dA}{dt} = \kappa m_A - \gamma A + \frac{1}{\sqrt{\Omega}}\left(\sqrt{\kappa m_A}\Gamma_1 - \sqrt{\gamma A}\Gamma_2\right),$$

in which Γ_5 and Γ_6 are the noise on the production reaction and decay reaction of mRNA, respectively. For the open loop system we have $F(A) = \alpha_0 p_{\text{tot}}$, while for the closed loop system we have the Hill function

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

The steady state for the open loop system is given by

$$m_e^o = \frac{\alpha_0 p_{\text{tot}}}{\delta}, \qquad A_e^o = \frac{\kappa\alpha_0 p_{\text{tot}}}{\delta\gamma}.$$

Considering Γ_6 as the input of interest, the linearization of the system at this equilibrium is given by

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

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

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

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

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

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

In order to make a fair comparison between the two systems, we let the steady states be the same. To do this, we 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 site.

The open loop and closed loop transfer functions are given by

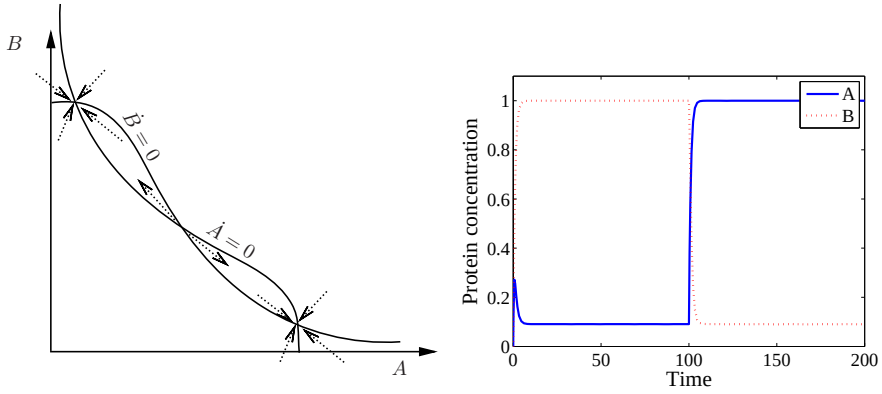
$$G^o_{A\Gamma_6}(s) = \frac{\kappa\sqrt{\delta m_e/\Omega}}{(s+\delta)(s+\gamma)},$$

and by

$$G^c_{A\Gamma_6}(s) = \frac{\kappa\sqrt{\delta m_e/\Omega}}{s^2 + s(\delta+\gamma) + \delta\gamma + \kappa g},$$

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 Γ_6 can be amplified at sufficiently high frequencies. Figure 6.2(b) shows the magnitude $M(\omega)$ of the corresponding frequency responses for both the open loop and the closed loop system.

It is clear that the presence of negative autoregulation 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 model, 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 towards the high frequency, as also experimentally reported [6].



(a) (b)

Figure 6.3: (a) 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. (b) 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 $\gamma = 1$. The inducers are such that $u_1 = 100$ for $t < 100$ and $u_1 = 0$ for $t \geq 100$, while $u_2 = 0$ for $t < 100$ and $u_2 = 100$ for $t \geq 100$.

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 6.1 [27]. We start by describing a simple model 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

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

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 $dB/dt = 0$ and letting $n \geq 2$, we obtain the nullclines shown in Figure 6.3a. The nullclines intersect at three points, which determine the steady states of this system. The stability of these steady states can be determined as follows.

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 deduces that the steady state at which $A = B$ is unstable while the other two are stable. Hence, this is a bistable system.

The system's trajectory converges to one steady state or the other depending on the initial condition. Specifically, a trajectory starting at an initial condition in

the region of attraction of steady state S, converges to this steady state. The 45 degree line divides the plane into the two regions of attraction of the stable steady states. Once the system's trajectory has converged to one of the two steady states, it cannot switch to the other unless an external stimulation is applied.

In the toggle switch by [27], external stimulations were added in the form of negative inducers for A and B. Specifically, 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. Hence, the system becomes

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

in which we have let $K_B = K(1+u_2/K_{d,2})$ and $K_A = K(1+u_1/K_{d,1})$ denote the effective K values of the Hill functions. We show in Figure 6.3b time traces for $A(t)$ and $B(t)$ when the inducer concentrations are changed. 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.

Note that the effect of the inducers in this model is that of changing the shape of the nullclines by changing the values of K_A and K_B. Specifically, high values of u_1 and $u_2 = 0$ will lead to increased values of K_A, which will shift the point of half-maximal value of the Hill function $\beta/(1+(A/K_A)^n)$ to the right. As a consequence, the nullclines will intersect at one point only, in which the value of B is high and the value of A is low. The opposite will occur when u_2 is high and $u_1 = 0$, leading to only one intersection point in which B is low and A is high.

6.4 The Repressilator

Elowitz and Leibler [23] 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

$$\begin{array}{lll} \frac{dm_A}{dt} = F_1(C) - \delta m_A & \frac{dm_B}{dt} = F_2(A) - \delta m_B & \frac{dm_C}{dt} = F_3(B) - \delta m_C \\ \frac{dA}{dt} = \kappa m_A - \gamma A & \frac{dB}{dt} = \kappa m_B - \gamma B & \frac{dC}{dt} = \kappa m_C - \gamma C, \end{array} \tag{6.2}$$

where we take

$$F_1(P) = F_2(P) = F_3(P) = \frac{\alpha}{1 + (P/K)^n}.$$

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 equilibrium is unstable, then the system admits a periodic solution. Therefore, to apply these results, 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. Letting $\beta = (\kappa/\delta)$, we obtain

$$A = \frac{\beta F_1(C)}{\gamma}, \qquad B = \frac{\beta F_2(A)}{\gamma}, \qquad C = \frac{\beta F_3(B)}{\gamma},$$

which combined together yield

$$A = \frac{\beta}{\gamma} F_1\left(\frac{\beta}{\gamma} F_3\left(\frac{\beta}{\gamma} F_2(A)\right)\right) =: g(A).$$

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

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

it follows that 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 [59]. In the present case, system (6.2) 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.

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

$$J = \begin{pmatrix} -\delta & 0 & 0 & 0 & 0 & F_1'(P) \\ \kappa & -\gamma & 0 & 0 & 0 & 0 \\ 0 & F_2'(P) & -\delta & 0 & 0 & 0 \\ 0 & 0 & \kappa & -\gamma & 0 & 0 \\ 0 & 0 & 0 & F_3'(P) & -\delta & 0 \\ 0 & 0 & 0 & 0 & \kappa & -\gamma \end{pmatrix},$$

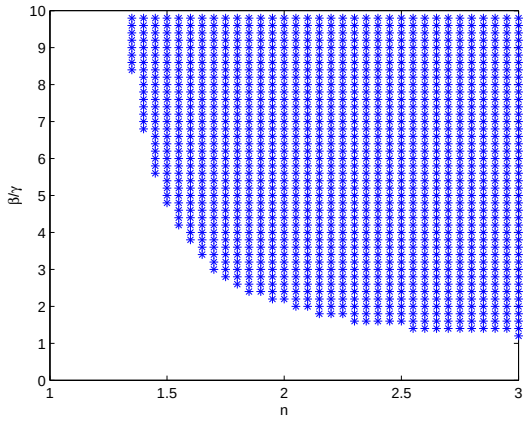


Figure 6.4: Space of parameters that give rise to oscillations for the repressilator in equation (6.2). Here, we have set $K = 1$ for simplicity.

whose characteristic polynomial is given by

$$\det(\lambda I - J) = (\lambda+\gamma)^3(\lambda+\delta)^3 - \kappa^3 \Pi_{i=1}^3 F_i'(P).$$

The roots of this characteristic polynomial are given by

$$(\lambda+\gamma)(\lambda+\delta) = s,$$

in which $s \in \{\kappa F'(P), (\kappa F'(P)/2)(1+i\sqrt{3}), (\kappa F'(P)/2)(1-i\sqrt{3})\}$. For being able to invoke Hastings Theorem to infer the existence of a periodic orbit, it is sufficient that one of the roots of the characteristic polynomial has positive real part. This is the case if

$$\kappa|F'(P)| > \gamma\delta, \qquad |F'(P)| = \alpha \frac{n(P^{n-1}/K^n)}{(1+(P/K)^n)^2},$$

in which P is the equilibrium value satisfying the equilibrium condition

$$P = \frac{\beta}{\gamma} \frac{\alpha}{1+(P/K)^n}.$$

One can plot the pair of values $(n, \beta/\gamma)$ for which the above two conditions are satisfied. This leads to the plot of Figure 6.4. When n increases, the existence of an unstable equilibrium point is guaranteed for larger ranges of β/γ. 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.

A similar result for the existence of a periodic solution can be obtained when two of the Hill functions are monotonically increasing and only one is monotonically decreasing:

$$F_1(P) = \frac{\alpha}{1+(P/K)^n}, \qquad F_2(P) = \frac{\alpha(P/K)^n}{1+(P/K)^n}, \qquad F_3(P) = \frac{\alpha(P/K)^n}{1+(P/K)^n}.$$

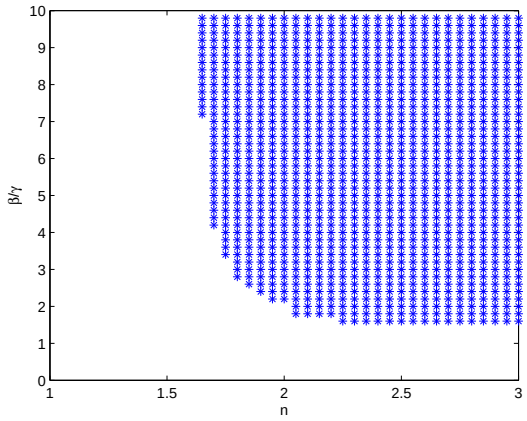


Figure 6.5: Space of parameters that give rise to oscillations for the repressilator (non-symmetric design). As the value of n is increased, the range of the other parameter for which a periodic cycle exists become larger. Here, we have set $K = 1$.

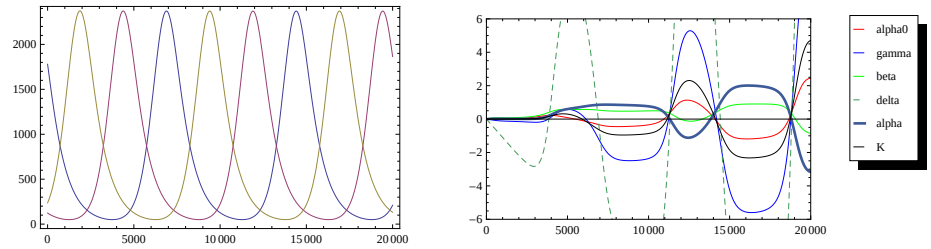
That is, two interactions are activations and one only is a repression. We refer to this as the "non-symmetric" design. Since the loop gain is still negative, there is only one equilibrium point. We can thus obtain the condition for oscillations again by establishing conditions on the parameters that guarantee that at least one root of the characteristic polynomial (6.4) has positive real part, that is,

$$\kappa(|F_1'(P_3)F_2'(P_1)F_3'(P_2)|)^{(1/3)} > \gamma\delta, \tag{6.3}$$

in which P_1, P_2, P_3 are the equilibrium values of A, B, and C. These satisfy:

$$P_2 = \frac{\beta}{\gamma}\frac{(P_1/K)^n}{1+(P_1/K)^n}, \qquad P_3 = \frac{\beta}{\gamma}\frac{(P_2/K)^n}{1+(P_2/K)^n}, \qquad P_1(1+(P_3/K)^n) = \frac{\beta}{\gamma}.$$

Using these expressions numerically and checking for each combination of the parameters $(n,\beta/\gamma)$ whether (6.3) is satisfied, we can plot the combinations of n and β/γ values that lead to an unstable equilibrium. This is shown in Figure 6.5. From this figure, we can deduce that the qualitative shape of the parameter space that leads to a limit cycle is the same in the repressilator and in the non-symmetric design. One can conclude that it is then possible to "over design" the circuit such that the parameters land in the filled region of the plots. In practice, values of the Hill coefficient n between one and two can be obtained by employing repressors that have cooperativity higher than or equal to two. There are plenty of such repressors, including those originally used in the repressilator design [23]. However, values of n greater than two may be hard to reach in practice. To overcome this problem, one can include more elements in the loop. In fact, it is possible to show that the value of n sufficient for obtaining an unstable equilibrium decreases when the number of elements in the loop is increased (see Exercises).



(a) (b)

Figure 6.6: (a) Repressilator proteins as functions of time. (b) Repressilator sensitivity plots. The most important parameters are the protein and mRNA decay rates γ and δ.

Example 6.1 (Repressilator Parameter Sensitivity)**.** In this example, we use the parameter sensitivity analysis tools of Chapter 3 to investigate the sensitivity of the protein concentrations to changes in the parameters. In this case, we model the repressilator Hill functions adding the basal expression rate as it was originally done in [23]:

$$F_1(P) = F_2(P) = F_3(P) = \frac{\alpha}{1+(P/K)^n} + \alpha_0.$$

Letting $x = (m_A, A, m_B, B, m_C, C)$ and $\theta = (\alpha_0, \delta, \kappa, \gamma, \alpha, K)$, we can compute the sensitivity $S_{x,\theta}$ along the limit cycle corresponding to nominal parameter vector θ_0 as illustrated in Chapter 3:

$$\frac{dS_{x,\theta}}{dt} = M(t,\theta_0)S_{x,\theta} + N(t,\theta_0),$$

where $M(t,\theta_0)$ and $N(t,\theta_0)$ are both periodic in time. If the dynamics of $S_{x,\theta}$ are stable then the resulting solutions will be periodic, showing how the dynamics around the limit cycle depend on the parameter values. The results are shown in Figure 6.6b, where we plot the steady state sensitivity of A as a function of time. We see, for example, that the limit cycle depends strongly on the protein degradation and dilution rate δ, indicating that changes in this value can lead to (relatively) large variations in the magnitude of the limit cycle.

∇

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:

$$\frac{dm_A}{dt} = -\delta_A m_A + F_1(A,B), \qquad \frac{dm_B}{dt} = -\delta_B m_B + F_2(A),$$
$$\frac{dA}{dt} = -\gamma_A A + \kappa_A m_A, \qquad \frac{dB}{dt} = -\gamma_B B + \kappa_B m_B,$$

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

$$F_1(A,B) = \frac{\alpha_A (A/K_A)^n + \alpha_{A0}}{1 + (A/K_A)^n + (B/K_B)^m}, \quad F_2(A) = \frac{\alpha_B (A/K_A)^n + \alpha_{B0}}{1 + (A/K_A)^n}.$$

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 (Section 3.4). Then, we analyze the four dimensional system and perform a bifurcation study.

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

$$\frac{dA}{dt} = f(A,B), \qquad \frac{dB}{dt} = g(A,B). \tag{6.4}$$

For simplifying notation, we assume $m = 1$ and $K_A = K_B = 1$.

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. Let $\overline{\alpha}_A = \alpha_A(\kappa_A/\delta_A)$, $\overline{\alpha}_B = \alpha_B(\kappa_B/\delta_B)$, $\overline{\alpha}_{A0} = \alpha_{A0}(\kappa_A/\delta_A)$, and $\overline{\alpha}_{B0} = \alpha_{B0}(\kappa_B/\delta_B)$. Then, $g(A,B) = 0$ leads to

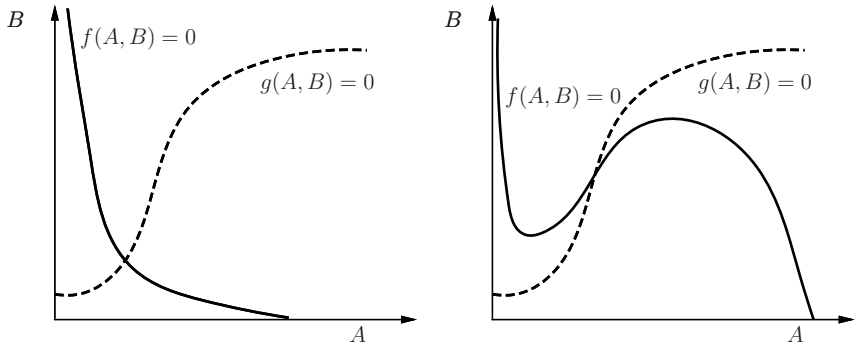
$$B = \frac{\overline{\alpha}_B A + \overline{\alpha}_{B0}}{(1+A)\gamma_B},$$

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

$$B = \frac{\overline{\alpha}_A A + \overline{\alpha}_{A0} - \gamma_A A(1+A)}{\gamma_A A},$$

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.



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

Figure 6.7: Nullclines for the two-dimensional system (6.4). (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.

This is given by the Jacobian matrix

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

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 Figure 6.7(a)) shows that

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

By the implicit function theorem (Chapter 3, Section 3.6), we further have that

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

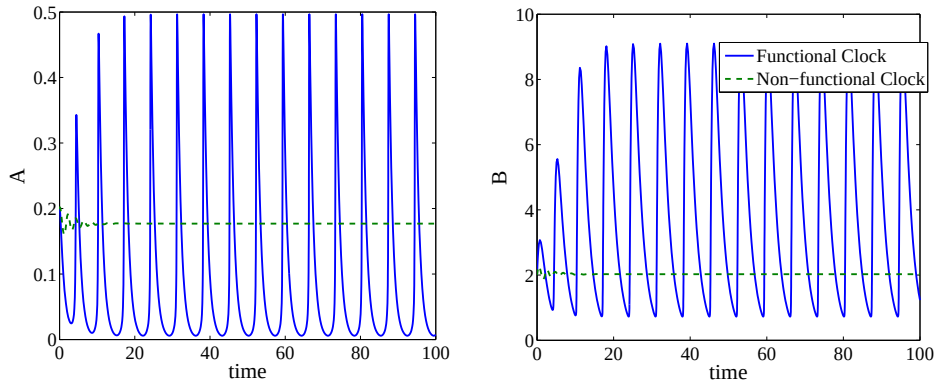
so that $\partial f/\partial A < 0$ because $\partial f/\partial B < 0$. As a consequence, we have that $\mathrm{tr}(J) < 0$ and hence the equilibrium point is either stable or a saddle.

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

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



(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 (6.4) for both a functional ($\mathrm{tr}(J) > 0$) and a non-functional ($\mathrm{tr}(J) < 0$) clock. The parameters in the simulation are $\delta_A = \delta_B = 1$, $\alpha_A = \alpha_B = 100$, $\alpha_{A0} = .04$, $\alpha_{B0} = .004$, $\gamma_A = 1$, $\kappa_A = \kappa_B = 1$, and $K_A = K_B = 1$. In the Functional Clock, $\gamma_B = 0.5$ whereas in the Non-Functional Clock, $\gamma_B = 1.5$. Parameters α_A and α_B were chosen to give about 500-2000 copies of protein per cell for activated promoters. Parameters α_{A0} and α_{B0} were chosen to give about 1-10 copies per cell for non-activated promoters.

so that $\det(J) > 0$. Hence, the ω-limit set (Chapter 3, Section 3.4) of any point in the plane is not necessarily a periodic orbit. It follows that 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 only one equilibrium point 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

$$\frac{\gamma_B}{\partial f_1/\partial A - \gamma_A} < 1.$$

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 γ_B is too large, the trace becomes negative and oscillations disappear.

Four-dimensional analysis

In order to deepen our understanding of the role of time scale separation between activator and repressor dynamics, we perform a time scale analysis employing the

bifurcation 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:

$$\frac{dm_A}{dt} = -(\delta_A/\epsilon)\, m_A + F_1(A,B), \qquad \frac{dm_B}{dt} = -(\delta_B/\epsilon)\, m_B + F_2(A),$$
$$\frac{dA}{dt} = \nu(-\gamma_A A + (\kappa_A/\epsilon)\, m_A), \qquad \frac{dB}{dt} = -\gamma_B B + (\kappa_B/\epsilon)\, m_B.$$

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 determines the relative time scale between the activator and the repressor dynamics. As ν increases, the activator dynamics become faster compared to the repressor dynamics. The parameter ϵ determines the relative time scale between the protein and mRNA dynamics. As ϵ becomes smaller, the mRNA dynamics become faster compared to protein dynamics and model (6.5) becomes close to the two-dimensional model (6.4), in which the mRNA dynamics are considered at the quasi-steady state. Thus, ϵ is a singular perturbation parameter. In particular, 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 as the two bifurcation parameters. The bifurcation analysis results are summarized by Figure 6.9. 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)

In Section 3.2, we described various mechanisms to obtain robustness to external perturbations. In particular, one such mechanism is provided by incoherent feedforward loops. Here, we describe an implementation that was proposed for making the steady state levels of protein expression robust to perturbations in DNA plasmid copy number [12]. In this implementation, the input u is the amount of DNA plasmid coding for both the intermediate regulator A and the output protein B. The intermediate regulator A represses through transcriptional repression the expression of the output protein B (Figure 6.10). The expectation is that the steady

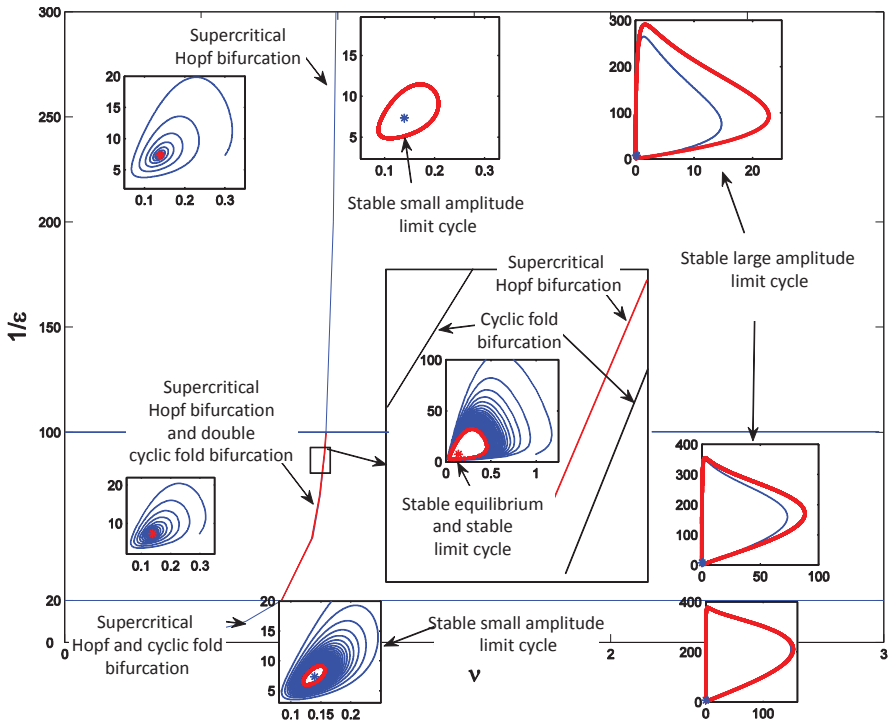


Figure 6.9: Design chart for the relaxation oscillator. We obtain sustained oscillations past the Hopf bifurcation point, for values of ν sufficiently large independently of the difference of time scales between the protein and the mRNA dynamics. We also notice that there are values of ν for which a 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 [55] and occurs for realistic values of the separation of time-scales between protein and mRNA 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.

state value of B is independent of the concentration u of the plasmid. That is, the concentration of B 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 that the mRNA dynamics are at the quasi-steady state. This model is given by

$$\frac{dA}{dt} = k_0 u - \gamma A, \qquad \frac{dB}{dt} = \frac{k_1 u}{1 + (A/K_\mathrm{d})} - \gamma B, \tag{6.5}$$

in which k_0 is the constitutive rate at which A is expressed and K_d is the dissociation

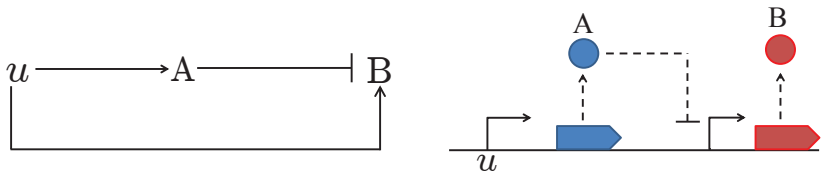


Figure 6.10: The incoherent feedforward motif with a possible implementation. The circuit is integrated on a DNA plasmid denoted u. Protein A is under the control of a constitutive promoter in the DNA plasmid u, while B is repressed by A. Protein B, in turn, is also expressed by a gene in the plasmid u. Hence B is also "activated" by u.

constant of the binding of A with the promoter. This implementation has been called the sniffer in Section 3.2. The steady state of the system is obtained by setting the time derivatives to zero and gives

$$A = \frac{k_0}{\gamma} u, \qquad B = \frac{k_1 u}{\gamma + k_0 u / K_\mathrm{d}}.$$

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 $B = k_1 K_\mathrm{d}/k_0$, 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 B. In practice, however, the value of K_d is non-zero, hence the adaptation is not perfect. We show in Figure 6.11 the behavior of the steady state of B 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 not modeled the cooperativity of the binding of protein A to the promoter. If A is LacI, for example, the cooperativity of binding is $n = 4$. We leave as an exercise to show that the adaptation behavior persists 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 ($1/K_\mathrm{d}$ in the Hill function) of a transcription factor to its site on the promoter can be weakened 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 transcription factors (combinatorial promoters) to implement regulation functions that take multiple inputs can be realized by combining the operator sites of several simple promoters [18].

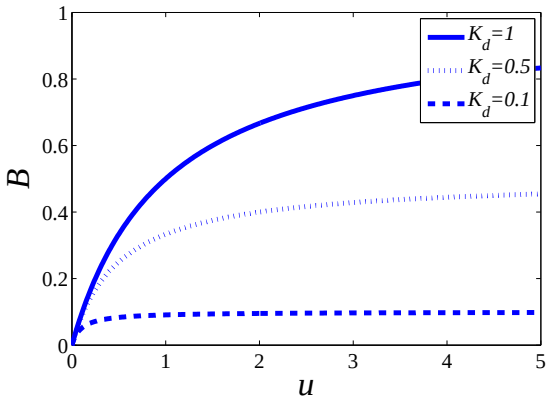


Figure 6.11: Behavior of the steady state value of B as a function of the input u.

Exercises

6.1 Consider the negatively autoregulated system:

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

Explore through linearization how increasing the Hill coefficient affects the response time of the system. Also, compare the results of the linearization analysis to the behavior of the nonlinear system obtained through simulation.

6.2 Consider the toggle switch:

$$\frac{dA}{dt} = \frac{\beta_A}{1+(B/K)^n} - \gamma A, \qquad \frac{dB}{dt} = \frac{\beta_B}{1+(A/K)^m} - \gamma B.$$

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).

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

6.3 Consider the repressilator model and the parameter space for oscillations provided in Figure 6.4. Determine how this parameter space changes if the value of K in the Hill function is changed.

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

6.5 Consider the oscillator design of Stricker et al. [86]. 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. Explore through simulation conditions for oscillations and compare the behavior of the reduced model to that of the original model.

6.6 Consider the feedforward circuit shown in Figure 6.10. Assume now to model cooperativity such that the model modifies to

$$\frac{dA}{dt} = k_0 u - \gamma A, \qquad \frac{dB}{dt} = \frac{k_1 u}{1 + (A/K_d)^4} - \gamma B.$$

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

Chapter 7
Interconnecting Components

In Chapter 2 and Chapter 6, we studied the behavior of simple biomolecular modules, such as oscillators, toggles, self repressing circuits, signal transduction and amplification systems, based on reduced order models. One natural step forward is to create larger and more complex systems by composing these modules together. In this chapter, we illustrate problems that need to be overcome when interconnecting components and propose a number of engineering solutions based on the feedback principles introduced in Chapter 3. Specifically, we explain how impedance-like effects arise at the interconnection between modules, which change the expected circuit behavior. These impedance problems appear in several other engineering domains, including electrical, mechanical, and hydraulic systems, and have been largely addressed by the respective engineering communities. In this chapter, we explain how similar engineering solutions can be employed in biomolecular systems to defeat impedance effects and guarantee "modular" interconnection of circuits. In Chapter 8, we further study loading of the cellular environment by synthetic circuits employing the same framework developed in this chapter.

7.1 Input/Output Modeling and the Modularity Assumption

The input/output modeling introduced in Chapter 1 and further developed in Chapter 3 has been employed so far to describe the behavior of various modules and subsystems. This input/output description of a system allows to connect systems together by setting the input u_2 of a downstream system equal to the output y_1 of

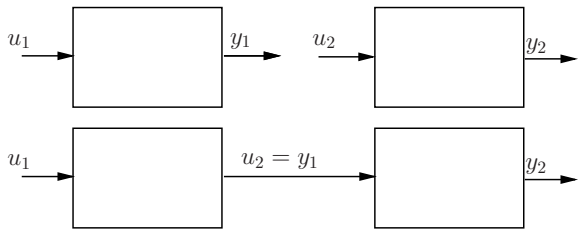


Figure 7.1: In the input/output modeling framework, systems are interconnected by statically assigning to the input of the downstream system the value of the output of the upstream system.

the upstream system (Figure 7.1) and has been extensively used in the previous chapters.

Each node of a gene circuit, such as those in Figure 6.1, has been modeled as an input/output system taking the concentrations of transcription factors as input and giving, through the processes of transcription and translation, the concentration of another transcription factor as an output. For example, node C in the repressilator has been modeled as a second order system that takes the concentration of transcription factor B as an input through the Hill function and gives transcription factor C as an output. This is of course not the only possible choice for delimiting a system. We could in fact let the mRNA or the RNA polymerase flowing along the DNA, called PoPS (polymerase per second) [25], play the role of input and output signals. Similarly, a signal transduction network is usually composed by protein covalent modification modules, which take a modifying enzyme (a kinase in the case of phosphorylation) as an input and gives the modified protein as an output. Accordingly, one of the models of the MAPK cascade considered in Section 2.5 is obtained by setting the value of the kinase concentration of a downstream cycle equal to the value of the concentration of the modified protein of the upstream cycle.

This input/output modeling framework is extremely useful because it allows to predict the behavior of an interconnected system from the behavior of the isolated modules. For example, the location and number of steady states in the toggle switch of Section 6.3 were predicted by intersecting the steady state input/output characteristics, determined by the Hill functions, of the isolated modules A and B. Similarly, the number of steady states in the repressilator was predicted by modularly composing the input/output steady state characteristics, again determined by the Hill functions, of the three modules composing the circuit.

For this input/output interconnection framework to reliably predict the behavior of connected modules, it is necessary that the input/output (dynamic) behavior of a system does not change upon interconnection to another system. We refer to the property by which a system input/output behavior does not change upon interconnection as *modularity*. All the designs and modeling described in the previous chapter assume that the modularity property holds. In this chapter, we question this assumption and investigate when modularity holds in gene and in signal transduction circuits. Further, we illustrate design methods, based on the techniques of Chapter 3, to create functionally modular systems.

7.2 Introduction to Retroactivity

The modularity assumption implies that when two modules are connected together, their behavior does not change because of the interconnection. However, a fundamental systems-engineering issue that arises when interconnecting subsystems is how the process of transmitting a signal to a “downstream” component affects the

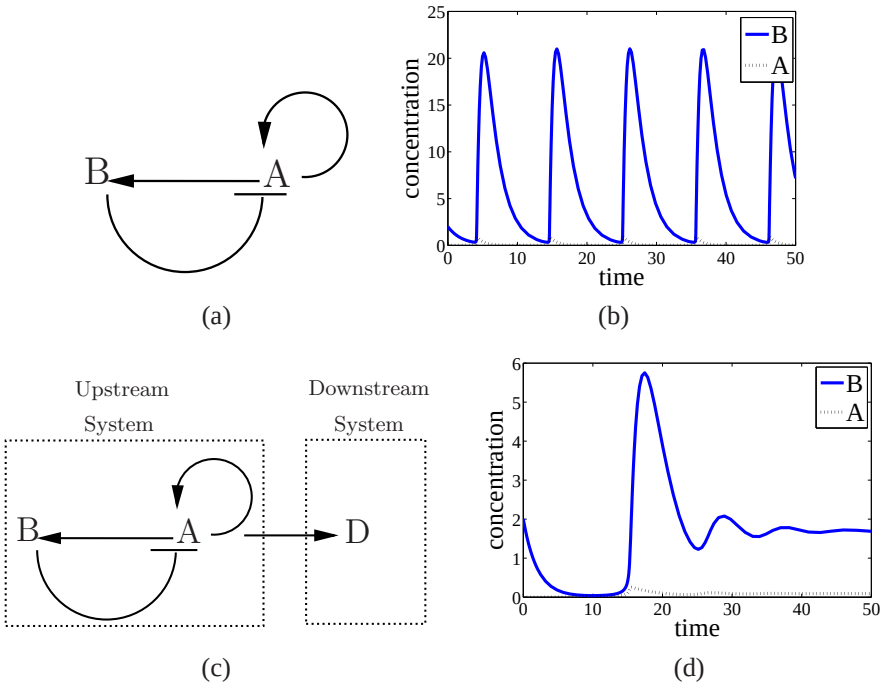


Figure 7.2: (a)-(b) show the activator-repressor clock topology and the time behavior of the activator and repressor concentrations. (c)-(d) show that when a load is connected to the clock, sustained oscillations disappear.

dynamic state of the sending component. This issue, the effect of "loads" on the output of a system, is well-understood in many engineering fields such as electrical engineering. It has often been pointed out that similar issues may arise for biological systems. These questions are especially delicate in design problems, such as those described in Chapter 6.

For example, consider a biomolecular clock, such as the activator-repressor clock introduced in Section 6.5 and shown in Figure 7.2a with simulations in Figure 7.2b. Assume that the activator protein concentration $A(t)$ is now used as a communicating species to synchronize or time a downstream system D (Figure 7.2c). From a systems/signals point of view, $A(t)$ becomes an *input* to the downstream system D. The terms "upstream" and "downstream" reflect the direction in which we think of signals as traveling, *from* the clock *to* the systems being synchronized. However, this is only an idealization because when A is taken as an input by the downstream system it binds to (and unbinds from) the promoter that controls the expression of D. These additional binding/unbinding reactions compete with the biochemical interactions that constitute the upstream clock and may therefore disrupt the operation of the clock itself (Figure **??**(d)). We call this "back-effect" *retroactivity* to extend the notion of impedance or loading to non-electrical systems and in partic-

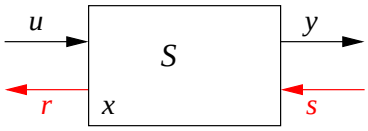


Figure 7.3: A system S input and output signals. The r and s signals denote signals originating by retroactivity upon interconnection [19].

ular to biomolecular systems. This phenomenon, while in principle may be used in an advantageous way by natural systems, can be deleterious when designing synthetic systems.

One possible approach to avoid disrupting the behavior of the clock is to introduce a gene coding for a new protein X, placed under the control of the same promoter as the gene for A, and using the concentration of X, which presumably mirrors that of A, to drive the downstream system. This approach, however, still has the problem that the behavior of the X concentration may be altered and even disrupted by the addition of downstream systems that drain X, as we shall see in the next section. The net result is that the downstream systems are not properly timed as X does not transmit the desired signal.

To model a system with retroactivity, we add to the input/output modeling framework used so far, an additional input, called s, to model any change that may occur upon interconnection with a downstream system. That is, s models the fact that whenever y is taken as an input to a downstream system the value of y may change, because of the physics of the interconnection. This phenomenon is also called in the physics literature "the observer effect", implying that no physical quantity can be measured without being altered by the measurement device. Similarly, we add a signal r as an additional output to model the fact that when a system is connected downstream of another one, it will send a signal upstream that will alter the dynamics of that system. More generally, we define a system S to have internal state x, two types of inputs, and two types of outputs: an input "u", an output "y" (as before), a *retroactivity to the input* "r", and a *retroactivity to the output* "s" (Figure 7.3). We will thus represent a system S by the equations

$$\frac{dx}{dt} = f(x,u,s), \qquad y = h(x,u,s), \qquad r = R(x,u,s), \tag{7.1}$$

where f, g, and R are arbitrary functions and the signals x, u, s, r, and y may be scalars or vectors. In such a formalism, we define the input/output model of the isolated system as the one in equation (7.1) without r in which we have also set $s = 0$.

Let S_i be a system with inputs u_i and s_i and with outputs y_i and r_i. Let S_1 and S_2 be two systems with disjoint sets of internal states. We define the interconnection of an upstream system S_1 with a downstream system S_2 by simply setting $y_1 = u_2$

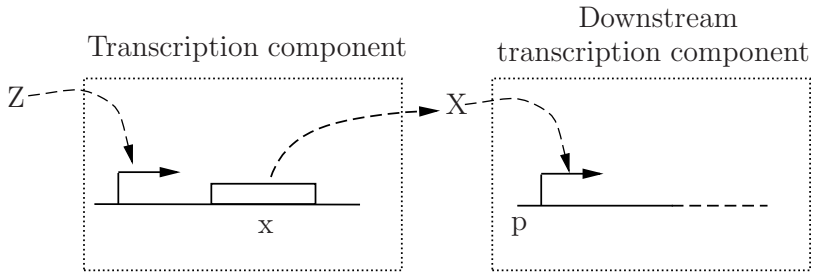


Figure 7.4: A transcription component takes as input u protein concentration Z and gives as output y protein concentration X. The downstream transcription component takes protein concentration X as its input.

and $s_1 = r_2$. For interconnecting two systems, we require that the two systems do not have internal states in common.

It is important to note that while retroactivity s is a back-action from the downstream system to the upstream one, it is conceptually different from feedback. In fact, retroactivity s is non-zero any time y is transmitted to the downstream system, that is, it is not possible to send signal y to the downstream system without retroactivity s. By contrast, feedback from the downstream system can be removed even when the upstream system sends signal y.

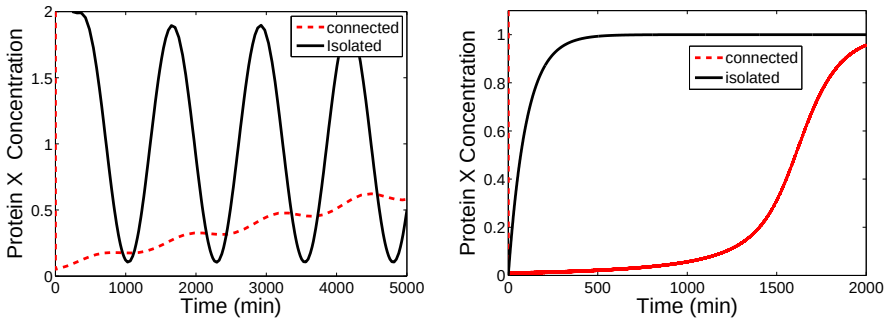
7.3 Retroactivity in Gene Circuits

In the previous section, we have introduced retroactivity as a general concept modeling the fact that when an upstream system is input/output connected to a downstream one, its behavior can change. In this section, we focus on gene circuits and show what form retroactivity takes and what its effects are.

Consider the interconnection of two transcription components illustrated in Figure 7.4. A transcription component is an input/output system that takes the transcription factor concentration Z as input and gives the transcription factor concentration X as output. The activity of the promoter controlling gene x depends on the amount of Z bound to the promoter. If $Z = Z(t)$, such an activity changes with time. To simplify notation, we denote it by $k(t)$. We assume here that the mRNA dynamics are at their quasi-steady state. The reader can verify that all the results hold unchanged when the mRNA dynamics are included (see exercises). We write the dynamics of X as

$$\frac{dX}{dt} = k(t) - \gamma X, \tag{7.2}$$

in which γ is the decay rate constant of the protein. We refer to equation (7.2) as the *isolated system dynamics*.



(a) (b)

Figure 7.5: The effect of retroactivity. The solid line represents $X(t)$ originating by equations (7.2), while the dashed line represents $X(t)$ obtained by equation (7.3). Both transient and permanent behaviors are different. Here, $k(t) = 0.01(1 + sin(\omega t))$ with $\omega = 0.005$ in (a) and $\omega = 0$ in (b), $k_{on} = 10$, $k_{off} = 10$, $\gamma = 0.01$, $p_{tot} = 100$, $X(0) = 5$. The choice of protein decay rate (in min^{-1}) corresponds to a half life of about one hour. The frequency of oscillations is chosen to have a period of about 12 times the protein half life in accordance to what is experimentally observed in the synthetic clock of [5].

Now, assume that X drives a downstream transcription module by binding to a promoter p with concentration p (Figure 7.4). The reversible binding reaction of X with p is given by

$$\mathrm{X} + \mathrm{p} \underset{k_{off}}{\overset{k_{on}}{\rightleftharpoons}} \mathrm{C}$$

in which C is the complex protein-promoter and k_{on} and k_{off} are the association and dissociation rate constants of protein X to promoter site p. Since the promoter is not subject to decay, its total concentration p_{tot} is conserved so that we can write $p + C = p_{tot}$. Therefore, the new dynamics of X are governed by the equations

$$\frac{dX}{dt} = k(t) - \gamma X + [k_{off}C - k_{on}(p_{tot} - C)X], \qquad \frac{dC}{dt} = -k_{off}C + k_{on}(p_{tot} - C)X. \quad (7.3)$$

We refer to this system as the *connected* system. Comparing the rate of change of X in the connected system to that in the isolated system (7.2), we notice the additional rate of change $[k_{off}C - k_{on}(p_{tot} - C)X]$ of X in the connected system. Hence, we have

$$s = [k_{off}C - k_{on}(p_{tot} - C)X],$$

and $s = 0$ when the system is isolated. We can interpret s as being a mass flow between the upstream and the downstream system, similar to a current in electrical circuits.

How large is the effect of retroactivity s on the dynamics of X and what are the biological parameters that affect it? We focus on the retroactivity to the output s as

we can analyze the effect of the retroactivity to the input r on the upstream system by simply analyzing the dynamics of Z in the presence of the promoter regulating the expression of x.

The effect of retroactivity s on the behavior of X can be very large (Figure 7.5). By looking at Figure 7.5, we notice that the effect of retroactivity is to "slow down" the dynamics of $X(t)$ as the response time to a step input increases and the response to a periodic signal appears attenuated and phase-shifted. We will come back to this more precisely in the next section.

These effects are undesirable in a number of situations in which we would like an upstream system to "drive" a downstream one as is the case, for example, when a biological oscillator has to time a number of downstream processes. If, due to the retroactivity, the output signal of the upstream process becomes too low and/or out of phase with the output signal of the isolated system (as in Figure 7.5), the coordination between the oscillator and the downstream processes will be lost. We next provide a procedure to obtain an operative quantification of the effect of the retroactivity on the dynamics of the upstream system.

Quantification of the retroactivity to the output

In this section, we provide a general approach to quantify the retroactivity to the output. To do so, we quantify the difference between the dynamics of X in the isolated system (7.2) and the dynamics of X in the connected system (7.3) by establishing conditions on the biological parameters that make the two dynamics close to each other. This is achieved by exploiting the difference of time scales between the protein production and decay processes and binding/unbinding reactions, mathematically described by $k_{\text{off}} \gg k(t), \gamma$. By virtue of this separation of time scales, we can approximate system (7.3) by a one dimensional system describing the evolution of X on the slow manifold (see Section 3.6).

To this end, note that (7.3) is not in standard singular perturbation form: while C is a fast variable, X is neither fast nor slow since its differential equation includes both fast and slow terms. To explicitly model the difference of time scales, we let $z = X + C$ be the total amount of protein X (bound and free) and re-write system (7.3) in the new variables (z, C). Letting $\epsilon = \gamma/k_{\text{off}}$, $K_d = k_{\text{off}}/k_{\text{on}}$, $k_{\text{off}} = \gamma/\epsilon$, and $k_{\text{on}} = \gamma/(\epsilon K_{\text{d}})$, system (7.3) can be re-written as

$$\frac{dz}{dt} = k(t) - \gamma(z - C), \qquad \epsilon\frac{dC}{dt} = -\gamma C + \frac{\gamma}{K_{\text{d}}}(p_{\text{tot}} - C)(z - C), \tag{7.4}$$

in which z is a slow variable. The reader can check as an exercise that the slow manifold of system (7.4) is locally exponentially stable (see Exercises).

We can obtain an approximation of the dynamics of X in the limit in which ϵ is very small, by setting $\epsilon = 0$. This leads to

$$-\gamma C + \frac{\gamma}{K_{\text{d}}}(p_{\text{tot}} - C)X = 0 \rightarrow C = g(X) \text{ with } g(X) = \frac{p_{\text{tot}}X}{X + K_{\text{d}}}.$$

Since $dy/dt = dX/dt + dC/dt$, we have that $dy/dt = dX/dt + (dg/dX)dX/dt$. This along with $dy/dt = k(t) - \gamma X$ lead to

$$\frac{dX}{dt} = (k(t) - \gamma X)\left(\frac{1}{1 + dg/dX}\right). \tag{7.5}$$

The difference between the dynamics in equation (7.5) (the connected system after a fast transient) and the dynamics in equation (7.2) (the isolated system) is zero when the term $\frac{dg(X)}{dX}$ in equation (7.5) is zero. We thus consider the term $\frac{dg(X)}{dX}$ as a quantification of the retroactivity s after a fast transient in the approximation in which $\epsilon \approx 0$. We can also interpret the term $\frac{dg(X)}{dX}$ as a percentage variation of the dynamics of the connected system with respect to the dynamics of the isolated system at the quasi-steady state. We next determine the physical meaning of such a term by calculating a more useful expression that is a function of key biochemical parameters.

By using the implicit function theorem, one can compute the following expression for $dg(X)/dX$:

$$\frac{dg(X)}{dX} = \frac{p_{\text{tot}}/K_{\text{d}}}{(X/K_{\text{d}} + 1)^2} =: \mathcal{R}(X). \tag{7.6}$$

The retroactivity measure $\mathcal{R}$ is low whenever the ratio $p_{\text{tot}}/K_{\text{d}}$, which can be seen as an effective load, is low. This is the case if the affinity of the binding sites p is small (K_{d} large) or if p_{tot} is low. Also, the retroactivity measure is dependent on X in a nonlinear fashion and it is such that it is maximal when X is the smallest. The expression of $\mathcal{R}(X)$ provides an operative quantification of retroactivity: such an expression can in fact be evaluated once the dissociation constant of X is known, the concentration of the binding sites p_{tot} is known, and X is also measured. From (7.5) and from (7.6), it follows that the rate of change of X in the connected system is smaller than that in the isolated system, that is, retroactivity slows down the dynamics of the transcription system. This has been also experimentally reported in [46].

Summarizing, the modularity assumption introduced in Section 7.1 holds only when the value of $\mathcal{R}(X)$ is small enough. As a consequence, the design of a simple circuit can assume modularity if the interconnections among the composing modules can be designed so that the value of $\mathcal{R}(X)$ is low. From a design point of view, low retroactivity can be obtained by either choosing low-affinity binding sites p or by making sure that the amounts of p is not too high compared to X. This can be guaranteed by placing the promoter sites p on low copy number plasmids or even on the chromosome (with copy number equal to 1). High copy number plasmids are expected to lead to non-negligible retroactivity effects on X.

In the presence of very low affinity and/or very low amount of promoter sites, the amount of complex C will be very low. As a consequence, the amplitude of the transmitted signal to downstream systems may also be very small and, as a consequence, noise may become a bottleneck. A better approach may be to design

insulation devices (as opposed to designing the interconnection for low retroactivity) to buffer systems from possibly large retroactivity as explained later in the chapter.

Effects of retroactivity on the frequency response

How do we explain the amplitude attenuation and phase shift due to retroactivity observed in Figure 7.5? In order to answer this question, we can linearize the system about its steady state and determine the effect of retroactivity on the frequency response. To this end, consider the input in the form $k(t) = \bar{k} + A_0 \sin(\omega t)$. Let $\bar{X} = \bar{k}/\gamma$ and $\bar{C} = p_{\text{tot}}\bar{X}/(\bar{X} + K_{\text{d}})$ be the equilibrium values corresponding to $\bar{k}$. The isolated system is already linear, so there is no need to perform linearization and the transfer function from k to X is given by

$$G^I_{Xk}(s) = \frac{1}{s+\gamma}.$$

For the connected system (7.5), let $(\bar{k}, \bar{X})$ denote the steady state, which is the same as for the isolated system, and let $\tilde{k} = k - \bar{k}$ and $x = X - \bar{X}$ denote small perturbations about this steady state. Then, the linearization of system (7.5) about $(\bar{k}, \bar{X})$ is given by (see Section 3.1):

$$\frac{dx}{dt} = (\tilde{k}(t) - \gamma x)\frac{1}{1 + (p_{\text{tot}}/K_{\text{d}})/(\bar{X}/K_{\text{d}} + 1)^2}.$$

Letting $\bar{R} := (p_{\text{tot}}/K_{\text{d}})/(\bar{X}/K_{\text{d}} + 1)^2$, we obtain the transfer function from $\tilde{k}$ to x of the connected system linearization as

$$G^C_{Xk} = \frac{1}{1+\bar{R}}\frac{1}{s + \gamma/(1+\bar{R})}.$$

Hence, we have the following result for the frequency response gain and phase shift:

$$M^I(\omega) = \frac{1}{\sqrt{\omega^2+\gamma^2}},\ \phi^I(\omega) = \tan^{-1}(-\omega/\gamma),$$

$$M^C(\omega) = \frac{1}{1+\bar{R}}\frac{1}{\sqrt{\omega^2 + \gamma^2/(1+R)^2}},\ \phi^C(\omega) = \tan^{-1}(-\omega(1+\bar{R})/\gamma),$$

from which one obtains that $M^I(0) = M^C(0)$ and, since $\bar{R} > 0$, the bandwidth of the connected system $\gamma/(1+\bar{R})$ is lower than that of the isolated system γ. As a consequence, we have that $M^I(\omega) > M^C(\omega)$ for all $\omega > 0$. Also, the phase shift of the connected system is larger than that of the isolated system. This explains why the plots of Figure 7.5 show a lag, an attenuation, and a phase shift in the response of the connected system.

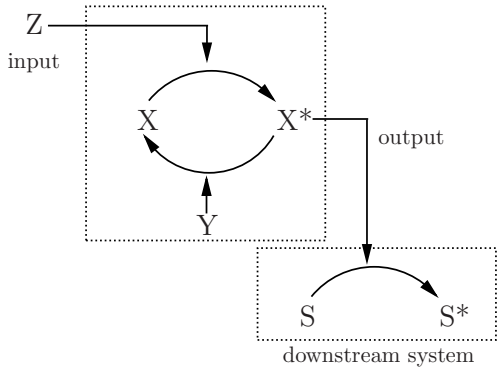


Figure 7.6: Covalent modification cycle with its input, output, and downstream system.

This fact further indicates that if the frequency of the input stimulation $k(t)$ is sufficiently lower than the bandwidth of the connected system $\gamma/(1+\bar{R})$, then the connected and isolated systems will respond similarly. Hence, the effects of retroactivity are tightly related to the time scale properties of the input signals and of the system, These effects will be negligible when the input stimulation is sufficiently slow (see exercises), and mitigation of retroactivity is required only when the frequency range of the signals of interest is larger than the connected system bandwidth $\gamma/(1+\bar{R})$.

7.4 Retroactivity in Signaling Systems

Signaling systems are circuits that take external stimuli as inputs and, through a sequence of biolmolecular reactions, transform them to useful signals that determine how cells respond to their environment. These systems are usually composed of covalent modification cycles (phosphorylation, methylation, urydylilation, etc.) connected in cascade fashion, in which each cycle has multiple downstream targets (or substrates). An example is the MAPK cascade, which we have analyzed in Section 2.5. Since covalent modification cycles always have downstream targets, such as DNA binding sites or other substrates, it is particularly important to understand whether and how retroactivity from these downstream systems affects the response of the upstream cycles to input stimulation. In this section, we study this question both for the steady state and dynamic response of a covalent modification cycle to its input (refer to Figure 7.6).

Steady state effects of retroactivity

We have seen in Section 2.4 that one important characteristic of signaling systems and, in particular, of covalent modification cycles, is the steady state characteristic or dose response curve. In particular, we showed in Section 2.4 that when the Michaelis-Menten constants are sufficiently small compared to the total protein amount, the cycle characteristics becomes ultrasensitive, a condition called zero-order ultrasensitivity. When the cycle is connected to its downstream targets, this steady state characteristics changes. In order to understand how this happens, we rewrite the reaction rates and the corresponding differential equation model for the covalent modification cycle of Section 2.4 adding the binding of X^* to its downstream target S. Referring to Figure 7.6, we have the following reactions:

$$Z + X \underset{d_1}{\overset{a_1}{\rightleftharpoons}} C_1 \xrightarrow{k_1} X^* + Z, \qquad Y + X^* \underset{d_2}{\overset{a_2}{\rightleftharpoons}} C_2 \xrightarrow{k_2} X + Y,$$

to which we add the binding reaction of X* with its substrates S:

$$X^* + S \underset{k_{\text{off}}}{\overset{k_{\text{on}}}{\rightleftharpoons}} C,$$

in which C is the complex of X* with S. In addition to this, we have the conservation laws $X_{\text{tot}} = X^* + X + C_1 + C_2 + C$, $Z + C_1 = Z_{\text{tot}}$, and $Y + C_2 = Y_{\text{tot}}$.

The rate equations governing the system are given by

$$\begin{aligned}
\frac{dC_1}{dt} &= a_1 XZ - (d_1 + k_1)C_1 \\
\frac{dX^*}{dt} &= -a_2 X^* Y + d_2 C_2 + k_1 C_1 - k_{\text{on}} S X^* + k_{\text{off}} C \\
\frac{dC_2}{dt} &= a_2 X^* Y - (d_2 + k_2)C_2 \\
\frac{dC}{dt} &= k_{\text{on}} X^* S - k_{\text{off}} C.
\end{aligned}$$

The input/output characteristics are found by solving this system for the equilibrium. In particular, by setting $dC_1/dt = 0$, $dC_2/dt = 0$, using that $Z = Z_{\text{tot}} - C_1$ and that $Y = Y_{\text{tot}} - C_2$, we obtain the familiar expressions for the complexes:

$$C_1 = \frac{Z_{\text{tot}} X}{K_1 + X}, \qquad C_2 = \frac{Y_{\text{tot}} X^*}{K_2 + X^*}, \qquad \text{with} \qquad K_1 = \frac{d_1 + k_1}{a_1} \qquad \text{and} \qquad K_2 = \frac{d_2 + k_2}{a_2}.$$

By setting $dX^*/dt + dC_2/dt + dC/dt = 0$, we obtain $k_1 C_2 = k_2 C_2$, which leads to

$$V_1 \frac{X}{K_1 + X} = V_2 \frac{X^*}{K_2 + X^*}, \qquad V_1 = k_1 Z_{\text{tot}} \qquad \text{and} \qquad V_2 = k_2 Y_{\text{tot}}. \tag{7.7}$$

By assuming that the substrate X_{tot} is in excess compared to the enzymes, we have that $C_1, C_2 \ll X_{\text{tot}}$ so that $X \approx X_{\text{tot}} - X^* - C$, in which (from setting $dC/dt = 0$) $C =$

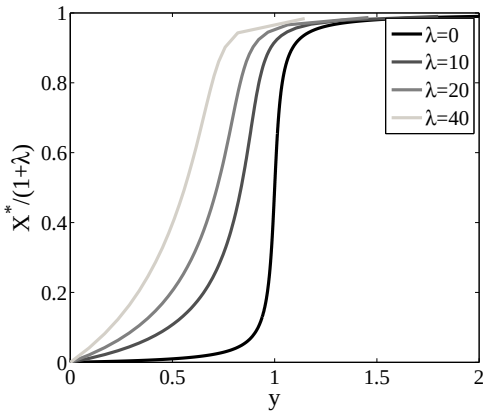


Figure 7.7: The addition of downstream target sites make the input/output characteristics more linear-like, that is, retroactivity makes a switch-like response into a more graded response.

X^*S/K_d with $K_d = k_{off}/k_{on}$, leading to $X \approx X_{tot} - X^*(1 + S/K_d)$. Calling $\lambda = S/K_d$, equation (7.7) finally leads to

$$y := \frac{V_1}{V_2} = \frac{X^*((K_1/1+\lambda) + ((X_{tot}/1+\lambda) - X^*))}{(K_2 + X^*)((X_{tot}/(1+\lambda)) - X^*)}. \tag{7.8}$$

Here, we can interpret λ as an effective load, which increases with the amount of targets of X^* but also with the affinity of these targets $(1/K_d)$.

We are interested in how the shape of the steady state characteristics of X^* as function of y change when the effective load λ is changed. As seen in Section 2.4, a way to quantify the sensitivity of the steady state characteristics is to calculate the response coefficient $R = y_{90}/y_{10}$. The maximal value of X^* obtained as $y \to \infty$ is given by $X_{tot}/(1+\lambda)$. Hence, from equation (7.8), we have that

$$y_{90} = \frac{(\bar{K_1} + 0.1)0.9}{(\bar{K_2}(1+\lambda) + 0.9)0.1}, \qquad y_{10} = \frac{(\bar{K_1} + 0.9)0.1}{(\bar{K_2}(1+\lambda) + 0.1)0.9},$$

$$\bar{K_1} := \frac{K_1}{X_{tot}}, \qquad K_2 = \frac{K_2}{X_{tot}},$$

so that

$$R = 81\frac{(\bar{K_1} + 0.1)(\bar{K_2}(1+\lambda) + 0.1}{(\bar{K_2}(1+\lambda) + 0.9)(\bar{K_1} + 0.9)}.$$

Comparing this expression with the one obtained in equation (2.25) for the isolated covalent modification cycle, we see that the net effect of the downstream target S is that of increasing the Michaelis-Menten constant K_2 by the factor $(1+\lambda)$. Hence, we should expect that with increasing load, the steady state characteristics should

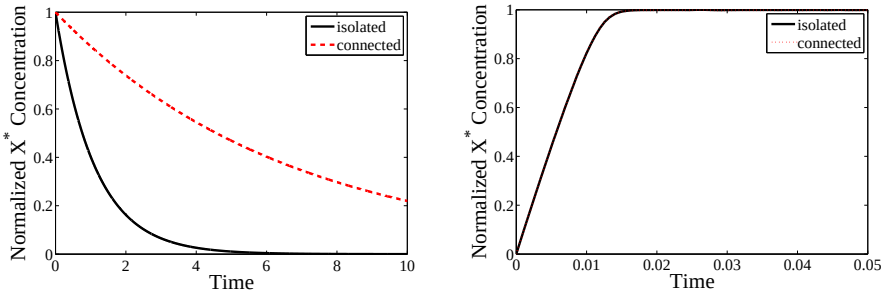


Figure 7.8: (a) Step response of the cycle in the presence of a positive step. The response time is not affected by the load. (b) Response to a negative step. The presence of the load makes the response slower. Here, $X_{\text{tot}} = 1$, $K_1 = K_2 = 0.1$, $k_1 = k_2 = 1$, and $\lambda = 5$.

be more linear-like. This is confirmed by the simulations shown in Figure 7.7 and it was also experimentally demonstrated in signal transduction circuits reconstituted *in vitro* [91].

One can check that R is a monotonically increasing function of λ. In particular, as λ increases, the value of R tends to $81(\bar{K}_1 + 0.1)/(\bar{K}_2 + 0.9)$, which, in turn, tends to 81 for $\bar{K}_1, \bar{K}_2 \to \infty$. When $\lambda = 0$, we recover the results of Section 2.4.

Dynamic effects of retroactivity

In order to understand the dynamic effects of retroactivity on the signaling module, we seek a one dimensional approximation of the X^* dynamics, which can be easily analyzed. To do so, we exploit time scale separation and apply singular perturbation analysis.

Specifically, we have that $d_i, k_{\text{off}} \gg k_1, k_2$, so we can choose as a small parameter $\epsilon = k_1/k_{\text{off}}$ and slow variable $w = X^* + C + C_2$. By setting $\epsilon = 0$, we obtain that $C_1 = Z_{\text{tot}}X/(K_1 + X)$, $C_2 = Y_{\text{tot}}X^*/(K_2 + X^*) =: g(X^*)$, and $C = \lambda X^*$, in which Z_{tot} is time-varying input signal. Hence, the dynamics of the slow variable w on the slow manifold are given by

$$\frac{dw}{dt} = k_1 \frac{Z_{\text{tot}}(t)X}{K_1 + X} - k_2 Y_{\text{tot}} \frac{X^*}{X^* + K_2}.$$

Using $dw/dt = dX^*/dt + dC/dt + dC_2/dt$, $dC/dt = \lambda dX^*/dt$, $dC_2/dt = \partial g/\partial X^* dX^*/dt$, and the conservation law $X = X_{\text{tot}} - X^*(1 + \lambda)$, we finally obtain the approximated X^* dynamics as

$$\frac{dX^*}{dt} = \frac{1}{1+\lambda}\left(k_1 \frac{Z_{\text{tot}}(t)(X_{\text{tot}} - X^*(1+\lambda))}{K_1 + (X_{\text{tot}} - X^*(1+\lambda))} - k_2 Y_{\text{tot}} \frac{X^*}{X^* + K_2}\right), \tag{7.9}$$

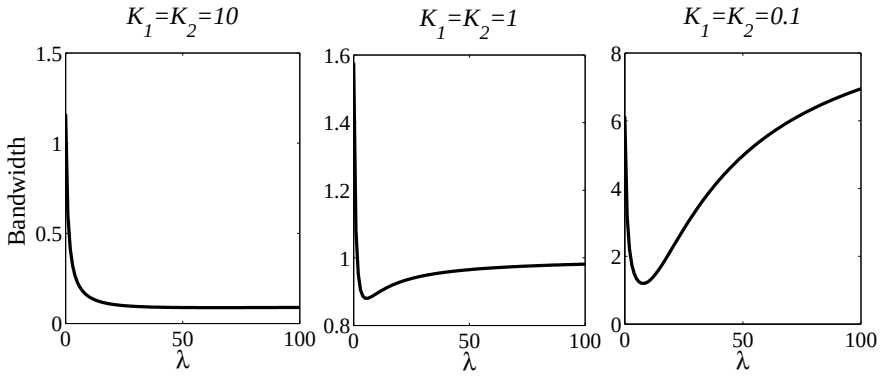


Figure 7.9: Behavior of the bandwidth as a function of the load for different values of the Michaelis-Menten constants K_1, K_2. Here, $X_{\text{tot}} = 1$.

where we have assumed that that $Y_{\text{tot}}/K_2 \ll S/K_{\text{d}}$, so that the effect of the binding dynamics of X* with Y (modeled by $\partial g/\partial X^*$) is negligible with respect to λ. The reader can verify this derivation as an exercise (see exercises).

From this expression, one can understand the effect of the load λ on the rise time and decay time in response to large step input stimuli Z_{tot}. For the decay time, one can assume an initial condition $X^*(0) \neq 0$ and $Z_{\text{tot}}(t) = 0$ for all t. In this case, we have that

$$\frac{dX^*}{dt} = -k_2 Y_{\text{tot}} \frac{X^*}{X^* + K_2} \frac{1}{1+\lambda},$$

from which, since $\lambda > 0$, it follows that the transient will be slower than when $\lambda = 0$ and hence that the system will have an increased decay time due to retroactivity. For the rise time, one can assume $Z_{\text{tot}} \approx \infty$ and $X^*(0) = 0$. In this case, at least initially we have that

$$(1+\lambda)\frac{dX^*}{dt} = \left(k_1 \frac{Z_{\text{tot}}(X_{\text{tot}} - X^*(1+\lambda))}{K_1 + (X_{\text{tot}} - X^*(1+\lambda))}\right),$$

which is the same expression for the isolated system in which X^* is scaled by $(1+\lambda)$. So, the rise time is not affected. The response of the cycle to positive and negative steps changes of the input stimulus Z are shown in Figure 7.8.

In order to understand how the bandwidth of the system is affected by retroactivity, we consider $Z_{\text{tot}}(t) = \bar{Z} + A_0 \sin(\omega t)$. Let $\bar{X}$ be the equilibrium of X^* corresponding to $\bar{Z}$. Let $z = Z_{\text{tot}} - \bar{Z}$ and $x = X^* - \bar{X}$ denote small perturbations about the equilibrium. The linearization of system (7.9) is given by

$$\frac{dx}{dt} = -a(\lambda)x + b(\lambda)z(t),$$

in which

$$a(\lambda) = \frac{1}{1+\lambda}\left(k_1\bar{Z}\frac{K_1(1+\lambda)}{(K_1+(X_{\text{tot}}-\bar{X}(1+\lambda)))^2} + k_2 Y_{\text{tot}}\frac{K_2}{(K_2+\bar{X})^2}\right)$$

and

$$b(\lambda) = \frac{k_1}{1+\lambda}\left(\frac{X_{\text{tot}}-\bar{X}(1+\lambda)}{K_1+(X_{\text{tot}}-\bar{X}(1+\lambda))}\right),$$

so that the bandwidth of the system is given by $\omega_B = a(\lambda)$.

Figure 7.9 shows the behavior of the bandwidth as a function of the load. When the isolated system static characteristics are linear-like ($K_1, K_2 \gg X_{\text{tot}}$), the bandwidth monotonically decreases with the load. By contrast, when the isolated system static characteristics are ultrasensitive ($K_1, K_2 \ll X_{\text{tot}}$), the bandwidth of the connected system can be larger than that of the isolated system for sufficiently large amounts of loads. In these conditions, one should expect that the response of the connected system becomes faster than that of the isolated system. These theoretical predictions have been experimentally validated in a covalent modification cycle reconstituted *in vitro* [48].

7.5 Insulation Devices: Retroactivity Attenuation

As explained in the previous section, it is not always possible or advantageous to design the downstream system so that it applies low retroactivity because, for example, the downstream system may already have been designed and optimized for other purposes. A better approach, in analogy to what is performed in electrical circuits, is to design a device to be placed between the upstream system (the oscillator, for example) and the downstream load so that the device output is not changed by the load and the device does not affect the behavior of the upstream system. That is, the output of the device should follow the prescribed behavior independently of any loading applied by a downstream system.

Specifically, consider a system S such as the one shown in Figure 7.3. We would like to design such a system such that

(a) the retroactivity r to the input is very small;

(b) the effect of the retroactivity s to the output on the internal dynamics of the system is very small independently of s itself (retroactivity attenuation).

Such a system is said to have the *insulation* property and will be called an insulation device. Indeed, such a system will not affect an upstream system because $r \approx 0$ and it will keep the same output signal y *independently* of any connected downstream system. Of course, other requirements may be important, such as the stability of the device and the speed of response.

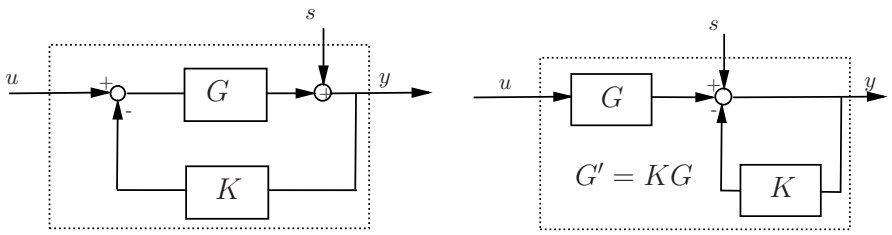


Figure 7.10: The diagram in (a) shows the basic high-gain feedback mechanism to attenuate the contribution of disturbance s on the output y. The diagram in (b) shows an alternative representation, which will be employed to design biological insulation devices.

Equation (7.6) quantifies the effect of retroactivity on the dynamics of X as a function of biochemical parameters. These parameters are the affinity of the binding site $1/K_{\mathrm{d}}$, the total concentration of such binding site p_{tot}, and the level of the signal $X(t)$. Therefore, to reduce retroactivity, we can choose parameters such that (7.6) is small. A sufficient condition is to choose K_{d} large (low affinity) and p_{tot} small, for example. Having small value of p_{tot} and/or low affinity implies that there is a small "flow" of protein X toward its target sites. Thus, we can say that a low retroactivity to the input is obtained when the "input flow" to the system is small. In the next sections, we focus on the retroactivity to the output, that is, on the retroactivity attenuation problem, and illustrate how the problem of designing a device that is robust to s can be formulated as a classical disturbance attenuation problem (Section 3.2). We provide two main design techniques to attenuate retroactivity: the first one (principle 1) is based on the idea of high-gain feedback (Section 3.2), while the second one uses time-scale separation and leverages the structure of the interconnection.

Attenuation of retroactivity to the output: Principle 1

The basic mechanism for retroactivity attenuation is based on the concept of disturbance attenuation through high-gain feedback presented in Section 3.2. In its simplest form, it can be illustrated by the diagram of Figure 7.10a, in which the retroactivity to the output s plays the same role as an additive disturbance. For large gains G, the effect of the retroactivity s to the output is negligible as the following simple computation shows. The output y is given by

$$y = G(u - Ky) + s,$$

which leads to

$$y = u\frac{G}{1+KG} + \frac{s}{1+KG}.$$

As G grows, y tends to u/K, which is independent of the retroactivity s.

Figure 7.10b illustrates an alternative representation of the diagram depicting high-gain feedback. This alternative depiction is particularly useful as it highlights that to attenuate retroactivity we need to (1) amplify the input of the system through a large gain and (2) apply a similarly large negative feedback on the output. The question of how to realize a large input amplification and a similarly large negative feedback on the output through biomolecular interactions is the subject of the next section. In what follows, we fist illustrate how this strategy also works for a dynamical system of the form of (7.5).

Consider the dynamics of the connected transcription system (7.5). Assume that we can apply a gain G to the input $k(t)$ and a negative feedback gain G' to X with $G' = KG$. This leads to the new differential equation for the connected system (7.5) given by

$$\frac{dX}{dt} = (Gk(t) - (G' + \gamma)X)(1 - d(t)), \tag{7.10}$$

in which we have defined $d(t) = \mathcal{R}(X)/(1 + \mathcal{R}(X))$. Since $d(t) < 1$, letting $G' = KG$, we can verify (see exercises) that as G grows $X(t)$ tends to $k(t)/K$ for both the connected system (7.10) and the isolated system

$$\frac{dX}{dt} = Gk(t) - (G' + \gamma)X. \tag{7.11}$$

Specifically, we have the following fact:

Proposition 7.1. *Consider the scalar system $\dot{x} = G(t)(k(t) - Kx)$ with $G(t) \geq G_0 > 0$ and $\dot{k}(t)$ bounded. Then, there are positive constants C_0 and C_1 such that*

$$|x(t) - k(t)/K| \leq C_0 e^{-G_0 K t} + \frac{C_1}{G_0}.$$

As a consequence, the solutions $X(t)$ of the connected and isolated systems tend to each other as G increases. Hence, the presence of the disturbance $d(t)$ will not significantly affect the time behavior of $X(t)$. It follows that the effect of retroactivity can be arbitrarily attenuated by increasing gains G and G'.

The next questions we address is how we can implement such amplification and feedback gains in a biomolecular system.

Biomolecular realizations of Principle 1

In this section, we illustrate two possible biomolecular implementations to obtain a large input amplification gain and a similarly large negative feedback on the output. Both implementations realize the negative feedback through enhanced degradation. The first design realizes amplification through transcription activation, while the second design uses phosphorylation.

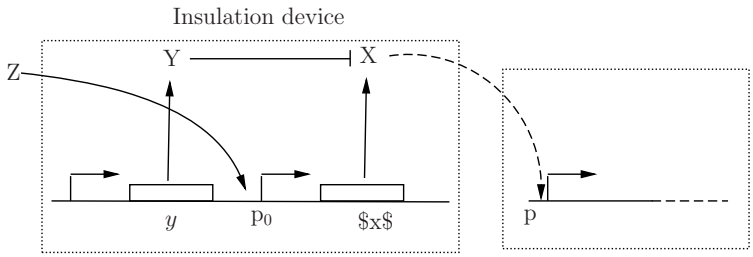


Figure 7.11: Design 1. The input $Z(t)$ is amplified by virtue of a strong promoter p_0. The negative feedback on the output X is obtained by enhancing its degradation through the protease Y.

Design 1: Amplification through transcription activation

This design is depicted in Figure 7.11. We implement a large amplification of the input signal $Z(t)$ by having Z be a transcription activator for protein X, such that the promoter p_0 controlling the expression of X is a strong, non-leaky promoter. The signal $Z(t)$ can be further amplified by increasing the strength of the ribosome binding site of gene x. The negative feedback mechanism on X relies on enhanced degradation of X. Since this must be large, one possible way to obtain an enhanced degradation for X is to have a specific protease, called Y, be expressed by a strong constitutive promoter.

To investigate whether such a design realizes a large amplification and a large negative feedback on X as needed to attenuate retroactivity to the output, we construct a model. The reaction of the protease Y with protein X is modeled as the two-step reaction $\mathrm{X}+\mathrm{Y} \underset{d}{\overset{a}{\rightleftharpoons}} \mathrm{W} \xrightarrow{\bar{k}} \mathrm{Y}$ (see Section 2.3). The input/output system model of the insulation device that takes Z as an input and gives X as an output is given by the following equations

$$\frac{dZ}{dt} = k(t) - \gamma_Z Z + \left[k_- \bar{C} - k_+ Z(p_{0,\mathrm{tot}} - \bar{C})\right] \tag{7.12}$$

$$\frac{d\bar{C}}{dt} = k_+ Z(p_{0,\mathrm{tot}} - \bar{C}) - k_- \bar{C} \tag{7.13}$$

$$\frac{dm_X}{dt} = G\bar{C} - \delta m_X \tag{7.14}$$

$$\frac{dW}{dt} = aXY - dW - \bar{k}W \tag{7.15}$$

$$\frac{dY}{dt} = -aYX + \bar{k}W + \alpha G - \gamma_Y Y + dW \tag{7.16}$$

$$\frac{dX}{dt} = \kappa m_X - aYX + dW - \gamma_X X + \left[k_{\mathrm{off}} C - k_{\mathrm{on}} X(p_{\mathrm{tot}} - C)\right] \tag{7.17}$$

$$\frac{dC}{dt} = -k_{\mathrm{off}} C + k_{\mathrm{on}} X(p_{\mathrm{tot}} - C), \tag{7.18}$$

in which we have assumed that the expression of gene z is controlled by a promoter with activity $k(t)$. In this system, we have denoted by k_+ and k_- the association and dissociation rates constants of Z with its promoter site p_0 in total concentration $p_{0,\text{tot}}$. Also, $\bar{C}$ is the complex of Z with such a promoter site. Here, m_X is the mRNA of X, and C is the complex of X bound to the downstream binding sites p with total concentration p_{tot}. The promoter controlling gene y has strength αG, for some constant α, and it has about the same strength as the promoter controlling x.

The terms in the square brackets in equation (7.12) represent the retroactivity r to the input of the insulation device in Figure 7.11. The terms in the square brackets in equation (7.17) represent the retroactivity s to the output of the insulation device. The dynamics of equations (7.12)–(7.18) without s describe the dynamics of X with no downstream system (isolated system).

Equations (7.12) and (7.13) determine the signal $\bar{C}(t)$ that is the input to equations (7.14)–(7.18). For the discussion regarding the attenuation of the effect of s, it is not relevant what the specific form of signal $\bar{C}(t)$ is. Let then $\bar{C}(t)$ be any bounded signal. Since equation (7.14) takes $\bar{C}(t)$ as an input, we will have that $m_X(t) = Gv(t)$, for a suitable signal $v(t)$. Let us assume for the sake of simplifying the analysis that the protease reaction is a one step reaction, that is, $\text{X} + \text{Y} \xrightarrow{\bar{k}'} \text{Y}$. Therefore, equation (7.16) simplifies to

$$\frac{dY}{dt} = \alpha G - \gamma_Y Y$$

and equation (7.17) simplifies to

$$\frac{dX}{dt} = \kappa m_X - \bar{k}' YX - \gamma_X X + k_{\text{off}} C - k_{\text{on}} X(p_{\text{tot}} - C).$$

If we further consider the protease to be at its equilibrium, we have that $Y(t) = \alpha G/\gamma_Y$.

As a consequence, the X dynamics become

$$\frac{dX}{dt} = \kappa G v(t) - (\bar{k}' \alpha G/\gamma_Y + \gamma_X)X + k_{\text{off}} C - k_{\text{on}} X(p_{\text{tot}} - C),$$

with C determined by equation (7.18). By using the same singular perturbation argument employed in the previous section, we obtain that the dynamics of X can be reduced to

$$\frac{dX}{dt} = (\kappa G v(t) - (\bar{k}' \alpha G/\gamma_Y + \gamma_X)X)(1 - d(t)), \tag{7.19}$$

in which $0 < d(t) < 1$ is the retroactivity term given by $\mathcal{R}(X)/(1 + \mathcal{R}(X))$. Then, as G increases, $X(t)$ becomes closer to the solution of the isolated system

$$\frac{dX}{dt} = \kappa G v(t) - (\bar{k}' \alpha G/\gamma_Y + \gamma_X)X,$$

as explained in the previous section be virtue of Proposition 7.1.

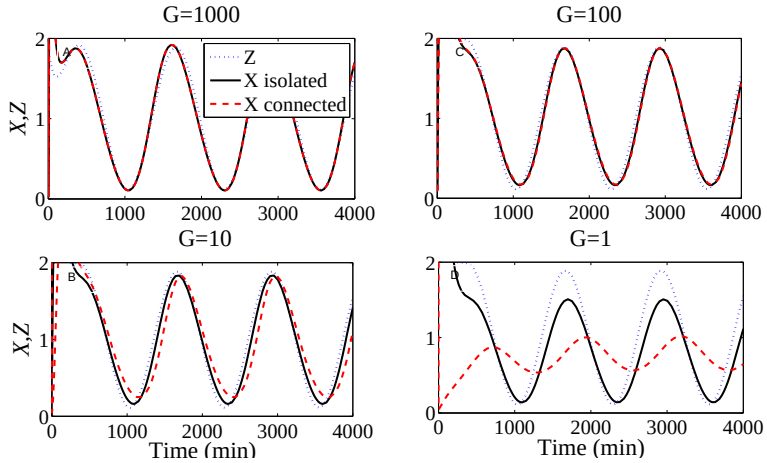


Figure 7.12: Design 1: results for different gains G. In all plots, $k(t) = 0.01(1 + sin(\omega t))$, $p_{\text{tot}} = 100$, $k_{\text{off}} = k_{\text{on}} = 10$, $\gamma_Z = 0.01 = \gamma_Y$, and $\omega = 0.005$. The parameter values are $\delta = 0.01$, $p_{0,\text{tot}} = 1$, $a = d = k' = 0.01$, $k_- = 200$, $k_+ = 10$, $\alpha = 0.1$, $\gamma_X = 0.1$, $\kappa = 0.1$, and $G = 1000, 100, 10, 1$. The retroactivity to the output is not well attenuated for values of the gain $G = 1$ and the attenuation capability begins to worsen for $G = 10$. Protein decay rates of 0.01min^{-1} correspond to a protein half life of about one hour. We consider a periodic forcing $k(t) = 0.01(1 + \sin(\omega t))$ with a period that is about 12 times the protein half life in accordance to what is experimentally observed in the synthetic clock of [5].

We now turn to the question of minimizing the retroactivity to the input r because its effect can alter the input signal $Z(t)$. In order to decrease r, we must guarantee that the retroactivity measure given in equation (7.6), in which we substitute Z in place of X, $p_{0,\text{tot}}$ in place of p_{tot}, and $K_{\text{d}} = k_+/k_-$ in place of K_{d}, is small. This is the case if $K_{\text{d}} \gg Z$ and $p_{0,\text{tot}}/K_{\text{d}} \ll 1$.

Simulation results for (7.12)–(7.18) are shown in Figure 7.12. For large gains ($G = 1000$, $G = 100$), the performance considerably improves compared to the case in which X was generated by a transcription component accepting Z as an input (Figure 7.5). For lower gains ($G = 10$, $G = 1$), the performance starts to degrade for $G = 10$ and becomes poor for $G = 1$. Since we can view G as the number of transcripts produced per unit time (one minute) per complex of protein Z bound to promoter p_0, values $G = 100, 1000$ may be difficult to realize *in vivo*, while the values $G = 10, 1$ could be more easily realized. However, the value of κ increases with the strength of the ribosome binding site and therefore the gain may be further increased by picking strong ribosme binding sites for x. The values of the parameters chosen in Figure 7.12 are such that $K_{\text{d}} \gg Z$ and $p_{0,\text{tot}} \ll K_{\text{d}}$. This is enough to guarantee that there is small retroactivity r to the input of the insulation device independently of the value of the gain G. The poorer performance of the device for $G = 1$ is therefore entirely due to poor attenuation of the retroactivity s to the output. To obtain a large negative feedback gain, we also require high expression

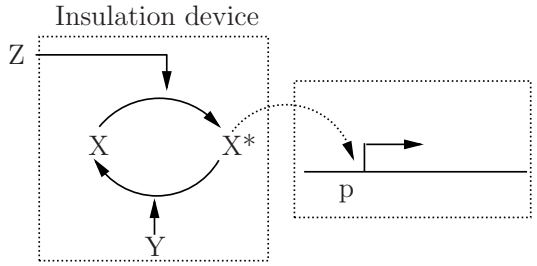


Figure 7.13: Phosphorylation cycle with a downstream DNA target. Amplification of Z occurs through the phosphorylation of substrate X. Negative feedback occurs through a phosphatase Y that converts the active form X* back to its inactive form X.

of the protease. It is therefore important that the protease is highly specific to its target X.

Design 2: Amplification through phosphorylation

In this design, the amplification gain G of Z is obtained by having Z be a kinase that phosphorylates a substrate X, which is available in abundance. The negative feedback gain G' on the phosphorylated protein X^* is obtained by having a phosphatase Y dephosphorylate the active protein X^*. Protein Y should also be available in abundance in the system. This implementation is depicted in Figure 7.13.

To illustrate what the key parameters are that enable retroactivity attenuation, we first consider a simplified model for the phosphorylation and dephosphorylation processes. This model will help in obtaining a conceptual understanding of what reactions are responsible in realizing the desired gains G and G'. The one step model that we consider is the same as considered in Chapter 2 (Exercise 2.8):

$$\mathrm{Z} + \mathrm{X} \xrightarrow{k_1} \mathrm{Z} + \mathrm{X}^*, \qquad \mathrm{Y} + \mathrm{X}^* \xrightarrow{k_2} \mathrm{Y} + \mathrm{X}.$$

We assume that there is an abundance of protein X and of phosphatase Y in the system and that these quantities are conserved. The conservation of X gives $X + X^* + C = X_{\mathrm{tot}}$, in which X is the inactive protein, X^* is the phosphorylated protein that binds to the downstream sites p, and C is the complex of the phosphorylated protein X^* bound to the promoter p. The X^* dynamics can be described by the following model

$$\frac{dX^*}{dt} = k_1 X_{\mathrm{tot}} Z(t)\left(1 - \frac{X^*}{X_{\mathrm{tot}}} - \left[\frac{C}{X_{\mathrm{tot}}}\right]\right) - k_2 Y X^* + [k_{\mathrm{off}} C - k_{\mathrm{on}} X^*(p_{\mathrm{tot}} - C)] \tag{7.20}$$

$$\frac{dC}{dt} = -k_{\mathrm{off}} C + k_{\mathrm{on}} X^*(p_{\mathrm{tot}} - C). \tag{7.21}$$

The terms in the square brackets represent the retroactivity s to the output of the insulation device of Figure 7.13. For a weakly activated pathway [38], $X^* \ll X_{\text{tot}}$. Also, if we assume that the total concentration of X is large compared to the concentration of the downstream binding sites, that is, $X_{\text{tot}} \gg p_{\text{tot}}$, equation (7.20) is approximatively equal to

$$\frac{dX^*}{dt} = k_1 X_{\text{tot}} Z(t) - k_2 Y X^* + k_{\text{off}} C - k_{\text{on}} X^* (p_{\text{tot}} - C).$$

Let $G = k_1 X_{\text{tot}}$ and $G' = k_2 Y$. Exploiting again the difference of time scales between the X^* dynamics and the C dynamics, the dynamics of X^* can be reduced to

$$\frac{dX^*}{dt} = (GZ(t) - G'X^*)(1 - d(t)),$$

in which $0 < d(t) < 1$ is the retroactivity term. Therefore, for G and G' large enough, $X^*(t)$ tends to the solution $X^*(t)$ of the isolated system $\frac{dX^*}{dt} = GZ(t) - G'X^*$, as explained before by virtue of Proposition 7.1. As a consequence, the effect of the retroactivity to the output s is attenuated by increasing the effective rates $k_1 X_{\text{tot}}$ and $k_2 Y$. That is, to obtain large input and negative feedback gains, one should have large phosphorylation/dephosphorylation rates and/or a large amount of protein X and phosphatase Y in the system. This reveals that the values of the phosphorylation/dephosphorylation rates cover an important role toward the retroactivity attenuation property of the module of Figure 7.13. The reader can verify through simulation how increasing the phosphatase and substrate amounts the effect of retroactivity can be attenuated (see exercises).

From a practical point of view, the effective rates can be increased by increasing the total amounts of X and Y. These amounts can be tuned, for example, by placing the *x* and *y* genes under the control of inducible promoters. Experiments performed on a covalent modification cycle reconstituted *in vitro*, showed that increasing these protein amounts is an effective means to attain retroactivity attenuation [48].

A design similar to the one illustrated can be proposed in which a phosphorylation cascade, such as the MAPK cascade, realizes the input amplification and an explicit feedback loop is added from the product of the cascade to its input [79]. The design presented here is simpler as it involves only one phosphorylation cycle and does not require any explicit feedback loop. In fact, a strong negative feedback can be realized by the action of the phosphatase that converts the active protein form X^* back to its inactive form X.

Attenuation of retroactivity to the output: Principle 2

In this section, we present a more general mechanism for retroactivity attenuation, which can be applied to systems of differential equations of arbitrary dimension. This will allow us to consider more complex and realistic models of the phospho-

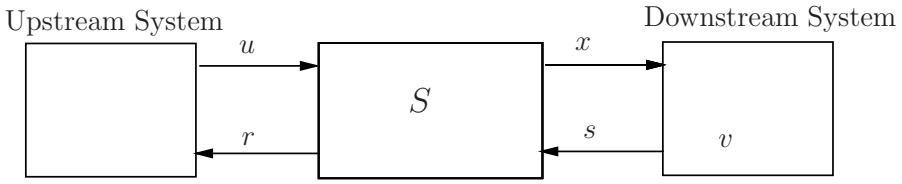


Figure 7.14: Interconnection of a device with input u and output x to a downstream system with internal state v applying retroactivity s.

rylation reactions and investigate the key parameters that control the retroactivity attenuation property.

For this purpose, consider Figure 7.14. We illustrate next how system S can attenuate retroactivity s by employing the principle of time scale separation. Specifically, when the internal dynamics of the system are much faster compared to the input u, the system immediately reaches its quasi-steady state with respect to the input. This quasi-steady state, in turn, is basically independent of s due to the interconnection structure between the systems. To illustrate this idea mathematically, consider the following simple structure in which (for simplicity) we assume that all variables are scalar:

$$\begin{aligned} \frac{du}{dt} &= f_0(u,t) + r(u,x) \\ \frac{dx}{dt} &= G f_1(x,u) + \bar{G} s(x,v) \\ \frac{dv}{dt} &= -\bar{G} s(x,v). \end{aligned} \tag{7.22}$$

Here let $G \gg 1$ to model the fact that the internal dynamics of the system are much faster than that of the input. Similarly, $\bar{G} \gg 1$ models the fact that the dynamics of the interconnection with downstream systems is also very fast. This is usually the case since the reactions in s are due to binding/unbinding reactions which are much faster than most of other biochemical processes, including gene expression and phosphorylation. The claim that we make about this system is the following.

> If $G \gg 1$ and the Jacobian of f_1 has eigenvalues with negative real part, then $x(t)$ is not affected by retroactivity s after a short initial transient, independently of the value of $\bar{G}$.

This result states that independently of the characteristics of the downstream system, system S can be tuned (by making G large enough) such that it attenuates the retroactivity to the output. To clarify why this would be the case, it is useful to rewrite system (7.22) in standard singular perturbation form by employing $\epsilon := 1/G$ as a small parameter and $\tilde{x} := x + v$ as the slow variable. Hence, it can be re-written

as

$$\begin{aligned} \frac{du}{dt} &= f_0(u,t) + r(u,x) \\ \epsilon\frac{d\tilde{x}}{dt} &= f_1(\tilde{x}-v,u) \\ \frac{dv}{dt} &= -\bar{G}s(\tilde{x}-v,v). \end{aligned} \tag{7.23}$$

Since $\partial f_1/\partial\tilde{x}$ has eigenvalues with negative real part, one can apply standard singular perturbation to show that after a very fast transient, the trajectories are attracted to the slow manifold given by $f_1(\tilde{x}-v,u)=0$. This is locally given by $x=g(u)$ solving $f_1(x,u)=0$. Hence, on the slow manifold we have that $x(t)=g(u(t))$, which is independent of the downstream system, that is, it is not affected by retroactivity.

The same result holds for a more general class of systems in which the variables u,x,v are vectors:

$$\begin{aligned} \frac{du}{dt} &= f_0(u,t) + r(u,x) \\ \frac{dx}{dt} &= G f_1(x,u) + \bar{G}As(x,v) \\ \frac{dv}{dt} &= -\bar{G}Bs(x,v) \end{aligned} \tag{7.24}$$

as long as there are matrices T and M such that $TA+MB=0$ and T is invertible. In fact, one can take the system to new coordinates $u,\tilde{x},v$ with $\tilde{x}=Tx+Mv$, in which the system will have the form (7.23).

Biomolecular realizations of Principle 2

We next consider possible biomolecular structures that realize Principle 2. Since this principle is based on a fast time scale of the device dynamics when compared to that of the device input, we focus on signaling systems, which are known to evolve on faster time scales than those of protein production and decay.

Design 1: Implementation through phosphorylation

We consider now a more realistic model for the phosphorylation and dephosphorylation reactions in a phosphorylation cycle than those considered in Section 7.5. In particular, we consider a two-step reaction model as seen in Section 2.4. According to this model, we have the following two reactions for phosphorylation and dephosphorylation:

$$\mathrm{X}+\mathrm{Z} \underset{d_1}{\overset{a_1}{\rightleftharpoons}} \mathrm{C}_1 \xrightarrow{k_1} \mathrm{X}^*+\mathrm{Z}, \qquad \mathrm{Y}+\mathrm{X}^* \underset{d_2}{\overset{a_2}{\rightleftharpoons}} \mathrm{C}_2 \xrightarrow{k_2} \mathrm{X}+\mathrm{Y}. \tag{7.25}$$

Additionally, we have the conservation equations $Y_{\text{tot}} = Y + C_2$, $X_{\text{tot}} = X + X^* + C_1 + C_2 + C$, because proteins X and Y are not degraded. Therefore, the differential equations modeling the system of Figure 7.13 become

$$\frac{dZ}{dt} = k(t) - \gamma Z\left[-a_1 Z X_{\text{tot}}(1 - \frac{X^*}{X_{\text{tot}}} - \frac{C_1}{X_{\text{tot}}} - \frac{C_2}{X_{\text{tot}}} - \left[\frac{C}{X_{\text{tot}}}\right]) + (d_1 + k_1)C_1\right] \quad (7.26)$$

$$\frac{dC_1}{dt} = -(d_1 + k_1)C_1 + a_1 Z X_{\text{tot}}(1 - \frac{X^*}{X_{\text{tot}}} - \frac{C_1}{X_{\text{tot}}} - \frac{C_2}{X_{\text{tot}}} - \left[\frac{C}{X_{\text{tot}}}\right]) \quad (7.27)$$

$$\frac{dC_2}{dt} = -(k_2 + d_2)C_2 + a_2 Y_{\text{tot}} X^*(1 - \frac{C_2}{Y_{\text{tot}}}) \quad (7.28)$$

$$\frac{dX^*}{dt} = k_1 C_1 + d_2 C_2 - a_2 Y_{\text{tot}} X^*(1 - \frac{C_2}{Y_{\text{tot}}}) + [k_{\text{off}} C - k_{\text{on}} X^*(p_{\text{tot}} - C)] \quad (7.29)$$

$$\frac{dC}{dt} = -k_{\text{off}} C + k_{\text{on}} X^*(p_{\text{tot}} - C), \quad (7.30)$$

in which the expression of Z is controlled by a promoter with activity $k(t)$. The terms in the large square bracket in equation (7.26) represent the retroactivity r to the input, while the terms in the square brackets of equations (7.27) and (7.29) represent the retroactivity s to the output.

We assume that $X_{\text{tot}} \gg p_{\text{tot}}$ so that in equations (7.26) and (7.27) we can neglect the term C/X_{tot} since $C < p_{\text{tot}}$. Phosphorylation and dephosphorylation reactions in equations (7.25) can occur at a much faster rate than protein production and decay processes (see Chapter 2). Choose X_{tot} and Y_{tot} to be sufficiently large, let $G = k_1 X_{\text{tot}}/\gamma$ and $\bar{G} = k_{\text{off}}/\gamma$. Then, we can re-write the system with $k_{\text{on}} = k_{\text{off}}/K_{\text{d}}$, $b_1 = a_1 X_{\text{tot}}/(\gamma G)$, $a_1 = a_2 Y_{\text{tot}}/(\gamma G)$, $b_2 = d_1/(\gamma G)$, $a_2 = d_2/(\gamma G)$, $c_i = k_i/(\gamma G)$, and $k_{\text{on}} = \bar{G}\gamma/K_{\text{d}}$. Letting $z = Z + C_1$ we obtain the system in the form

$$\begin{aligned}
\frac{dz}{dt} &= k(t) - \gamma(z - C_1) \\
\frac{dC_1}{dt} &= G\left(-\gamma(b_2 + c_1)C_1 + \gamma b_1(z - C_1)\left(1 - \frac{X^*}{X_{\text{tot}}} - \frac{C_1}{X_{\text{tot}}} - \frac{C_2}{X_{\text{tot}}})\right)\right) \\
\frac{dC_2}{dt} &= G\left(-\gamma(c_2 + a_2)C_2 + \gamma a_1 X^*\left(1 - \frac{C_2}{Y_{\text{tot}}}\right)\right) \\
\frac{dX^*}{dt} &= G\left(\gamma c_1 C_1 + \gamma a_2 C_2 - \gamma a_1 X^*\left(1 - \frac{C_2}{Y_{\text{tot}}}\right)\right) + \bar{G}(\gamma C - \gamma/K_{\text{d}}(p_{\text{tot}} - C)X^*) \\
\frac{dC}{dt} &= -\bar{G}(\gamma C - \gamma/K_{\text{d}}(p_{\text{tot}} - C)X^*),
\end{aligned} \quad (7.31)$$

which is in the form of system (7.24) with $u = z$, $x = (C_1, C_2, X^*)$, and $v = C$, in which one can choose T as the 3 by 3 identity matrix and $M = (0\ 0\ 1)'$. Hence, this system, for G sufficiently larger than 1 attenuates the effect of the retroactivity to the output s. For G to be large, one has to require that $k_1 X_{\text{tot}}$ is sufficiently

large and that $a_2 Y_{\text{tot}}$ is also comparatively large. These are the same design requirements obtained in the previous section based on the one-step reaction model of the enzymatic reactions.

In order to understand the effect of retroactivity to the input on the Z dynamics, one can consider the reduced system describing the dynamics on the time scale of Z. To this end, let $K_{m,1} = (d_1 + k_1)/a_1$ and $K_{m,2} = (d_2 + k_2)/a_2$ represent the Michaelis-Menten constants of the forward and backward enzymatic reactions, let $G = 1/\epsilon$ in (7.31), and take ϵ to the left-hand side. Setting $\epsilon = 0$ in the third and fourth equations of (7.31) the following relationships can be obtained:

$$C_1 = g_1(X^*) = \frac{(X^* Y_{\text{tot}} k_2)/(K_{m,2} k_1)}{1 + X^*/K_{m,2}}, \;\; C_2 = g_2(X^*) = \frac{(X^* Y_{\text{tot}})/K_{m,2}}{1 + X^*/K_{m,2}}. \tag{7.32}$$

Using expressions (7.32) in the second of equations (7.31) with $\epsilon = 0$ leads to

$$g_1(X^*)\left(b_2 + c_1 + \frac{b_1 Z}{X_{\text{tot}}}\right) = b_1 Z\left(1 - \frac{X^*}{X_{\text{tot}}} - \frac{g_2(X^*)}{X_{\text{tot}}}\right). \tag{7.33}$$

Assuming for simplicity that $X^* \ll K_{m,2}$, we obtain that $g_1(X^*) \approx X^* Y_{\text{tot}} k_2/K_{m,2} k_1$ and that $g_2(X^*) \approx X^*/K_{m,2} Y_{\text{tot}}$. As a consequence of these simplifications, equation (7.33) leads to

$$X^*(Z) = \frac{b_1 Z}{\frac{b_1 Z}{X_{\text{tot}}}(1 + Y_{\text{tot}}/K_{m,2} + (Y_{\text{tot}} k_2)/(K_{m,2} k_1)) + \frac{Y_{\text{tot}} k_2}{K_{m,2} k_1}(b_2 + c_1)}.$$

In order not to have distortion from Z to X^*, we require that

$$Z \ll \frac{Y_{\text{tot}} \frac{k_2}{k_1} \frac{K_m}{K_{m,2}}}{1 + \frac{Y_{\text{tot}}}{K_{m,2}} + \frac{Y_{\text{tot}}}{K_{m,2}} \frac{k_2}{k_1}}, \tag{7.34}$$

so that $X^*(Z) \approx Z X_{\text{tot}} K_{m,2} k_1 / Y_{\text{tot}} K_{m,1} k_2$ and therefore we have a linear relationship between X^* and Z with gain from Z to X^* given by $X_{\text{tot}} K_{m,2} k_1 / Y_{\text{tot}} K_{m,1} k_2$. In order not to have attenuation from Z to X^* we require that the gain is greater than or equal to one, that is,

$$\text{input/output gain} \approx \frac{X_{\text{tot}} K_{m,2} k_1}{Y_{\text{tot}} K_{m,1} k_2} \geq 1. \tag{7.35}$$

Requirements (7.34), (7.35) and $X^* \ll K_{m,2}$ are enough to guarantee that we do not have nonlinear distortion between Z and X^* and that X^* is not attenuated with respect to Z. In order to guarantee that the retroactivity r to the input is sufficiently small, we need to quantify the retroactivity effect on the Z dynamics due to the binding of Z with X. To achieve this, we proceed as in Section 7.3 by computing

the Z dynamics on the slow manifold, which gives a good approximation of the dynamics of Z if $\epsilon \approx 0$. These dynamics are given by

$$\frac{dZ}{dt} = (k(t) - \gamma Z)\left(1 - \frac{dg_1}{dX^*}\frac{dX^*}{dz}\right),$$

in which $\frac{dg_1}{dX^*}\frac{dX^*}{dz}$ measures the effect of the retroactivity r to the input on the Z dynamics. Direct computation of $\frac{dg_1}{dX^*}$ and of $\frac{dX^*}{dz}$ along with $X^* \ll K_{m,2}$ and with (7.34) leads to $\frac{dg_1}{dX^*}\frac{dX^*}{dz} \approx X_{\text{tot}}/K_{m,1}$, so that in order to have small retroactivity to the input, we require that

$$\frac{X_{\text{tot}}}{K_{m,1}} \ll 1. \tag{7.36}$$

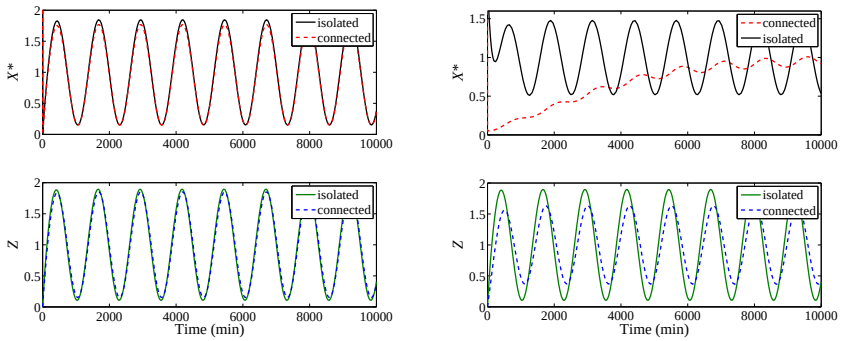
Hence, a design trade-off appears: X_{tot} should be sufficiently large to provide a gain G large enough to attenuate the retroactivity to the output. Yet, X_{tot} should be small enough compared to $K_{m,1}$ so to apply minimal retroactivity to the input.

Concluding, for having attenuation of the effect of the retroactivity to the output s, we require that the time scale of the phosphorylation/dephosphorylation reactions is much faster than the production and decay processes of Z (the input to the insulation device) and that $X_{\text{tot}} \gg p_{\text{tot}}$, that is, the total amount of protein X is in abundance compared to the downstream binding sites p. To obtain also a small effect of the retroactivity to the input, we require that $K_{m,1} \gg X_{\text{tot}}$. This is satisfied if, for example, kinase Z has low affinity to binding with X. To keep the input/output gain between Z and X^* close to one (from equation (7.35)), one can choose $X_{\text{tot}} = Y_{\text{tot}}$, and equal coefficients for the phosphorylation and dephosphorylation reactions, that is, $K_{m,1} = K_{m,2}$ and $k_1 = k_2$.

System in equations (7.26–7.30) was simulated with and without the downstream binding sites p, that is, with and without, respectively, the terms in the small box of equation (7.26) and in the boxes in equations (7.29) and (7.27). This is performed to highlight the effect of the retroactivity to the output s on the dynamics of X^*. The simulations validate our theoretical study that indicates that when $X_{\text{tot}} \gg p_{\text{tot}}$ and the time scales of phosphorylation/dephosphorylation are much faster than the time scale of decay and production of the protein Z, the retroactivity to the output s is very well attenuated (Figure 7.15a). Similarly, the time behavior of Z was simulated with and without the terms in the large box in equation (7.26), that is, with and without X to which Z binds, to verify whether the insulation device exhibits retroactivity to the input r.

In particular, the accordance of the behaviors of $Z(t)$ with and without its downstream binding sites on X (Figure 7.15a), indicates that there is no substantial retroactivity to the input r generated by the insulation device. This is obtained because $X_{\text{tot}} \ll K_{m,1}$ as indicated in equation (7.36), in which $1/K_m$ can be interpreted as the affinity of the binding of X to Z.

Our simulation study also indicates that a faster time scale of the phosphorylation/dephosphorylation reactions is necessary, even for high values of X_{tot} and



(a) Fast cycle (b) Slow cycle

Figure 7.15: (a) Performance with fast phosphorylation cycle. Simulation results for system in equations (7.26–7.30). In all plots, $p_{\text{tot}} = 100$, $k_{\text{off}} = k_{\text{on}} = 10$, $\gamma = 0.01$, $k(t) = 0.01(1 + sin(\omega t))$, and $\omega = 0.005$. In subplots A and B, $k_1 = k_2 = 50$, $a_2 = a_1 = 0.01$, $d_1 = d_2 = 10$, and $Y_{\text{tot}} = X_{\text{tot}} = 1500$. In the upper plot, the isolated system is without downstream binding sites p and the connected system is with binding sites p. The small error shows that the effect of the retroactivity to the output s is attenuated very well. In the lower plot, the isolated system stands for the case in which Z does not have X to bind to, while the connected system stands for the case in which Z binds to substrate X ($X_{\text{tot}} = 1500$). The small error confirms a small retroactivity to the input r. (b) Performance with a slow phosphorylation cycle. Phosphorylation and dephosphorylation rates are slower than the ones in (a), that is, $k_1 = k_2 = 0.01$, while the other parameters are left the same, that is, $d_2 = d_1 = 10$, $a_2 = a_1 = 0.01$, and $Y_{\text{tot}} = X_{\text{tot}} = 1500$.

Y_{tot}, to maintain perfect attenuation of the retroactivity to the output s and small retroactivity to the output r. In fact, slowing down the time scale of phosphorylation and dephosphorylation, the system looses its insulation property (Figure 7.15b). In particular, the attenuation of the effect of the retroactivity to the output s is lost because there is not enough separation of time scales between the Z dynamics and the internal device dynamics. The device also displays a non negligible amount of retroactivity to the input because the condition $K_m \ll X_{\text{tot}}$ is not satisfied anymore.

Design 2: Realization through phosphotransfer

Let X be a transcription factor in its inactive form and let X^* be the same transcription factor once it has been activated by the addition of a phosphate group. Let Z^* be a phosphate donor, that is, a protein that can transfer its phosphate group to the acceptor X. The standard phosphotransfer reactions (see Section 2.4) can be modeled according to the two-step reaction model

$$Z^* + X \underset{k_2}{\overset{k_1}{\rightleftharpoons}} C_1 \underset{k_4}{\overset{k_3}{\rightleftharpoons}} X^* + Z,$$

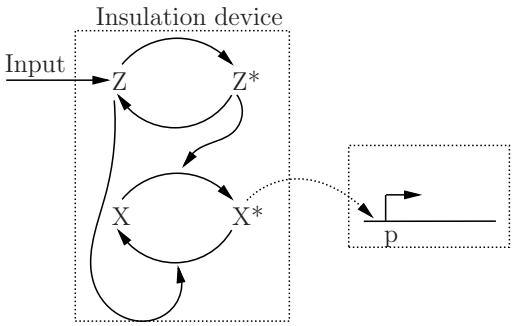
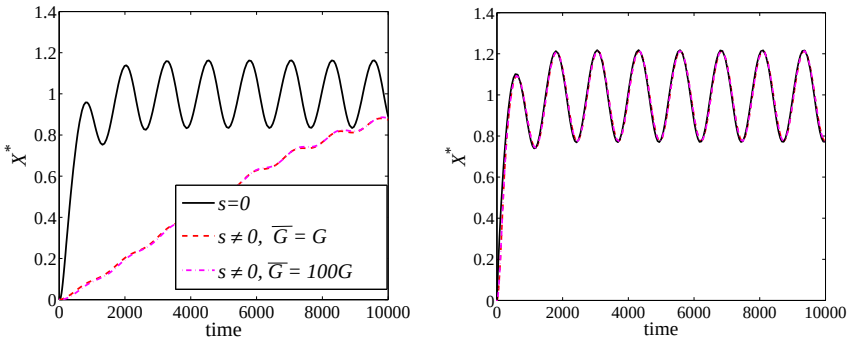


Figure 7.16: System S is a phosphotransfer system. The output X* activates transcription through the reversible binding of X* to downstream DNA promoter sites p.

in which C_1 is the complex of Z bound to X bound to the phosphate group. Additionally, protein Z can be phosphorylated and protein X^* dephosphorylated by other phosphotransfer interactions. These reactions are modeled as one-step reactions depending only on the concentrations of Z and X^*, that is, $Z \xrightarrow{\pi_1} Z^*$, $X^* \xrightarrow{\pi_2} X$. Protein X is assumed to be conserved in the system, that is, $X_{\text{tot}} = X + C_1 + X^* + C$. We assume that protein Z is produced with time-varying production rate $k(t)$ and decays with rate γ. The active transcription factor X^* binds to downstream DNA binding sites p with total concentration p_{tot} to activate transcription through the reversible reaction $p + X^* \underset{k_{\text{off}}}{\overset{k_{\text{on}}}{\rightleftharpoons}} C$. Since the total amount of p is conserved, we also have that $C + p = p_{\text{tot}}$. The ODE model corresponding to this system is thus given by the equations

$$\begin{aligned}
\frac{dZ}{dt} &= k(t) - \gamma Z + k_3 C_1 - k_4 X^* Z - \pi_1 Z \\
\frac{dC_1}{dt} &= k_1 X_{\text{tot}} \left(1 - \frac{X^*}{X_{\text{tot}}} - \frac{C_1}{X_{\text{tot}}} - \left[\frac{C}{X_{\text{tot}}}\right]\right) Z^* - k_3 C_1 - k_2 C_1 + k_4 X^* Z \\
\frac{dZ^*}{dt} &= \pi_1 Z + k_2 C_1 - k_1 X_{\text{tot}} \left(1 - \frac{X^*}{X_{\text{tot}}} - \frac{C_1}{X_{\text{tot}}} - \left[\frac{C}{X_{\text{tot}}}\right]\right) Z^* \\
\frac{dX^*}{dt} &= k_3 C_1 - k_4 X^* Z + [k_{\text{off}} C - k_{\text{on}} X^* (p_{\text{tot}} - C)] - \pi_2 X^* \\
\frac{dC}{dt} &= k_{\text{on}} X^* (p_{\text{tot}} - C) - k_{\text{off}} C.
\end{aligned} \tag{7.37}$$

Since phosphotransfer reactions are faster than protein production and decay, define $G := X_{\text{tot}} k_1 / \gamma$ so that $\bar{k}_1 := X_{\text{tot}} k_1 / G = \gamma$, $\bar{k}_2 := k_2 / G$, $\bar{k}_3 := k_3 / G$, $\bar{k}_4 := k_4 / G$, $\bar{\pi}_1 := \pi_1 / G$, $\bar{\pi}_2 := \pi_2 / G$ are of the same order of $k(t)$ and γ. Similarly, the process of protein binding and unbinding to promoter sites is much faster than protein production and decay. Let $\bar{G} := k_{\text{off}} / \gamma$ and $K_{\text{d}} := k_{\text{off}} / k_{\text{on}}$. Assuming also that $p_{\text{tot}} \ll$



(a) $G = 1$ (b) $G = 100$

Figure 7.17: Output response of the phosphotransfer system with a periodic signal $k(t) = \gamma(1+0.5\sin\omega t)$. The parameters are given by $\gamma = 0.01$, $X_{\text{tot}} = 5000$, $k_1 = k_2 = k_3 = k_4 = \pi_1 = \pi_2 = 0.01G$ in which $G = 1$ (a), and $G = 100$ (b). The downstream system parameters are given by $K_{\text{d}} = 1$ and $k_{\text{off}} = 0.01\bar{G}$, in which $\bar{G}$ assumes the values indicated on the legend. The isolated system ($s = 0$) corresponds to $p_{\text{tot}} = 0$ while the connected system ($s \neq 0$) corresponds to $p_{\text{tot}} = 100$.

X_{tot}, we have that $C \ll X_{\text{tot}}$ so that system (7.37) can be rewritten as

$$\begin{aligned}
\frac{dZ}{dt} &= k(t) - \gamma Z + G\left(\bar{k}_3 C_1 - \bar{k}_4 Z X^* - \bar{\pi}_1 Z\right)\\
\frac{dC_1}{dt} &= G\left(\bar{k}_1\left(1 - \frac{X^*}{X_{\text{tot}}} - \frac{C_1}{X_{\text{tot}}}\right)Z^* - \bar{k}_3 C_1 - \bar{k}_2 C_1 + \bar{k}_4 X^* Z\right)\\
\frac{dZ^*}{dt} &= G\left(\bar{\pi}_1 Z + \bar{k}_2 C_1 - \bar{k}_1\left(1 - \frac{X^*}{X_{\text{tot}}} - \frac{C_1}{X_{\text{tot}}}\right)Z^*\right)\\
\frac{dX^*}{dt} &= G\left(\bar{k}_3 C_1 - \bar{k}_4 X^* Z - \bar{\pi}_2 X^*\right) - \bar{G}\left(\frac{\gamma}{K_{\text{d}}} X^*(p_{\text{tot}} - C) + \gamma C\right)\\
\frac{dC}{dt} &= \bar{G}\left(\frac{\gamma}{K_{\text{d}}} X^*(p_{\text{tot}} - C) - \gamma C\right).
\end{aligned} \tag{7.38}$$

Taking $T = \mathbb{I}_{3\times 3}$, the 3 by 3 identity matrix, and $M = (0,0,1)^T$, the coordinate transformation $\tilde{x} = Tx + Mv$ brings the system to the form of system (7.24) with $u = Z$, $x = (C_1, Z^*, X^*)$, and $v = C$.

Figure 7.17a shows that, for a periodic input $k(t)$, the system with low value for G suffers from retroactivity to the output. However, for a large value of G (Figure 7.17b), the permanent behavior of the connected system becomes similar to that of the isolated system, whether $G \gg \bar{G}$, $G = \bar{G}$ or $G \ll \bar{G}$. This confirms the theoretical result that, independently of the order of magnitude of $\bar{G}$, the system can arbitrarily attenuate retroactivity for large enough G.

Exercises

7.1 Include in the study of retroactivity in transcription systems the mRNA dynamics and demonstrate how/whether the results change. Specifically, consider the following model of a connected transcription system

$$\begin{aligned}\frac{m_X}{dt} &= k(t) - \delta m_X \\ \frac{dX}{dt} &= \kappa m_X - \gamma X + [k_{\text{off}} C - k_{\text{on}}(p_{\text{tot}} - C)X], \\ \frac{dC}{dt} &= -k_{\text{off}} C + k_{\text{on}}(p_{\text{tot}} - C)X,\end{aligned}$$

7.2 Consider the system in standard singular perturbation form, in which $\epsilon \ll 1$. Demonstrate that the slow manifold is locally exponentially stable.

$$\frac{dz}{dt} = k(t) - \gamma(z - C), \qquad \epsilon\frac{dC}{dt} = -\gamma C + \frac{\gamma}{k_d}(p_{\text{tot}} - C)(z - C).$$

7.3 The characterization of retroactivity effects in a transcription module was based on the following model of the interconnection:

$$\begin{aligned}\frac{dX}{dt} &= k(t) - \gamma X + [k_{\text{off}} C - k_{\text{on}}(p_{\text{tot}} - C)X], \\ \frac{dC}{dt} &= -k_{\text{off}} C + k_{\text{on}}(p_{\text{tot}} - C)X,\end{aligned}$$

in which it was implicitly assumed that the complex C does not dilute. This is often a fair assumption. However, depending on the experimental conditions, a more appropriate model may include dilution for the complex C. In this case, the model modifies to

$$\begin{aligned}\frac{dX}{dt} &= k(t) - (\mu + \bar{\gamma})X + [k_{\text{off}} C - k_{\text{on}}(p_{\text{tot}} - C)X], \\ \frac{dC}{dt} &= -k_{\text{off}} C + k_{\text{on}}(p_{\text{tot}} - C)X - \mu C,\end{aligned}$$

in which μ represents decay due to dilution and $\bar{\gamma}$ represents protein degradation. Employ singular perturbation to determine the reduced X dynamics and the effects of retroactivity in this case. Is the steady state characteristic of the transcription module affected by retroactivity? How?

7.4 In this problem, we are going to study the frequency-dependent effects of retroactivity in gene circuits through simulation to validate the findings obtained

through linearization in Section 7.3. In particular, consider the model of a connected transcription component (7.3). Consider the parameters provided in Figure 7.5 and simulate the system with input $k(t) = \gamma(1 + sin(\omega t))$ with $\omega = 0.005$. Then, decrease and increase the frequency progressively and make a frequency/amplitude plot for both connected and isolated systems. Increase γ and re-do the frequency/amplitude plot. Comment on the retroactivity effects that you observe.

7.5 Consider the negatively autoregulated gene illustrated in Section 6.2. Instead of modeling negative autoregulation using the Hill function, explicitly model the binding of A with its own promoter. In this case, the formed complex C will be transcriptionally inactive (see Section 2.3). Explore through simulation how the response of the system without negative regulation compares to that with negative regulation when the copy number of the A gene is increased and the unrepressed expression rate β is decreased.

7.6 We have illustrated that the expression of the point of half-maximal induction in a covalent modification cycle is affected by the effective load λ as follows:

$$y_{50} = \frac{\bar{K}_1 + 0.5}{\bar{K}_2(1+\lambda) + 0.5}.$$

Study the behavior of this quantity when the effective load λ is changed.

7.7 Show how equation (7.9) is derived in Section 7.4.

7.8 Demonstrate through a mathematical proof that in the following system

$$\frac{dX}{dt} = G\,(k(t) - KX)\,(1 - d(t)),$$

in which $d(t) < 1$, we have that $X(t) - k(t)/K$ becomes smaller as G is increased.

7.9 Consider the one-step reaction model of the phosphorylation cycle with downstream binding sites given in (7.21). Simulate the system and determine how the behavior of the connected system compares to that of the isolated system when the amounts of substrate and phosphatase X_{tot} and Y_{tot} are increased.

7.10 Consider the activator-repressor clock described in Section 6.5 and take the parameter values of Figure 6.8 that result in a limit cycle. Then, assume that the activator A connects to another transcription circuit through the reversible binding of A with operator sites p to form activator-operator complex C: $\text{A} + \text{p} \underset{k_{off}}{\overset{k_{on}}{\rightleftharpoons}} \text{C}$ (connected clock). Answer the following questions:

(i) Simulate the connected clock and vary the total amount of p, that is, p_{tot}. Explore how this affects the behavior of the clock.

(ii) Give a mathematical explanation of the phenomenon you saw in (i). To do so, use singular perturbation to approximate the dynamics of the clock with downstream binding on the slow manifold (here, $k_{\text{on}}, k_{\text{off}} \gg \gamma_A, \gamma_B$).

(iii) Assume now that A does not bind to sites p, while the repressor B does. Take the parameter values of Figure 6.8 that result in a stable equilibrium. Explore how increasing p_{tot} affects the clock trajectories.

Chapter 8
Design Tradeoffs

In this chapter, we describe a couple of design tradeoffs arising from the interaction between synthetic circuits and the host organism. We specifically focus on two issues. The first issue is concerned with the effects of competition for shared cellular resources on circuits' behavior. In particular, circuits (endogenous and exogenous) share a number of cellular resources, such as RNA polymerase, ribosomes, ATP, nucleotides, etc. The insertion (induction) of synthetic circuits in the cell environment increases (changes) the demand for these resources, with possibly undesired repercussions on the functioning of the circuits themselves. Independent circuits may become actually coupled when they share common resources that are not in overabundence. This fact leads to constraints among the concentrations of proteins in synthetic circuits, which should be accounted for in the design phase. The second issue we consider, is the effect of biological noise on the design of devices requiring high gains. Specifically, we illustrate possible design tradeoffs between retroactivity attenuation, requiring high gains, and noise amplification, which emerge due to the intrinsic noise of biomolecular reactions.

8.1 Competition for Shared Cellular Resources

Exogenous circuits, just like the endogenous ones, use cellular resources, such as ribosomes, RNA polymerase, and ATP, that are shared among all the circuitry of the cell. From a systems and signals point of view, these interactions can be depicted as in Figure 8.1. The cell endogenous circuitry system produces resources as output and exogenous circuits take these resources as inputs. As a consequence, as seen in Chapter 7, there is retroactivity from the exogenous circuits to the engoneous circuitry system. This retroactivity creates indirect coupling between the exogenous circuits and can, in principle, lead to undesired crosstalk. In this chapter, we study the effect of the retroactivity from the synthetic circuits to shared resources in the cellular environment by focusing on the effect on availability of RNA polymerase and ribosomes, for simplicity. We then study the consequence of this retroactivity, illustrating how the behavior of unconnected circuits becomes coupled. These effects are significant for any resource whose availability is not in substantial excess compared to the demand of exogeous circuits.

In order to illustrate the problem, we consider the simple system shown in Figure 8.2, in which two modules, a constitutively expressed gene (Module 1) and a

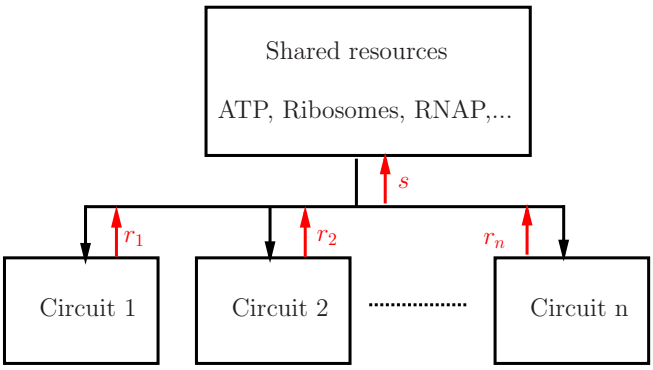


Figure 8.1: The cell environment provides resources to synthetic circuits, such as RNA polymerase, ribosomes, ATP, nucleotides, proteases, etc. These resources can be viewed as an "output" of the cell endogenous circuitry and an input to the exogenous circuits. Circuit *i* takes these resources as input and, as a consequence, it causes a retroactivity r_i to its input. Hence, the endogenous circuitry system has a retroactivity to the output *s* that encompasses all the retroactivities applied by the exogenous circuits.

gene activated by a transcription activator A (Module 2), are both present in the cellular environment. In theory, Module 2 should respond to changes in the activator A concentration, while Module 1, having a constitutively active promoter, should display a constant expression level that is independent of the activator A concentration. Experimental results, however, indicate that this is not the case: Module 1's output protein P_1 concentration also responds to changes in the activator A concentration. In particular, as the activator A concentration is increased, the concentration of protein P_1 can substantially decrease. This fact can be qualitatively explained by the following reasoning. When A is added, RNAP can bind to DNA promoter D_2 and start transcription, so that the free available RNAP decreases as some is bound on DNA D_2. Transcription of Module 2 generates mRNA and hence ribosomes will have more ribosome binding sites to which they can bind, so that less ribosomes will be free and available for other reactions. It follows that the addition of activator A leads to an overall decrease of the free RNAP and ribosomes that can take part in the transcription and translation reactions of Module 1. The net effect is that less P_1 protein will be produced.

The extent of this effect will depend on the overall availability of the shared resources and whether they are regulated. It is known that RNAP and ribosomes are internally regulated by a combination of feedback interactions [47, 56]. This, of course, may help compensating for changes in the demand of these resources.

In this chapter, we illustrate how this effect can be mathematically explained by explicitly accounting for the usage of RNAP and ribosomes in the transcription and translation models of the circuits. To simplify the mathematical analysis and to gather analytical understanding of the key parameters at the basis of this phe-

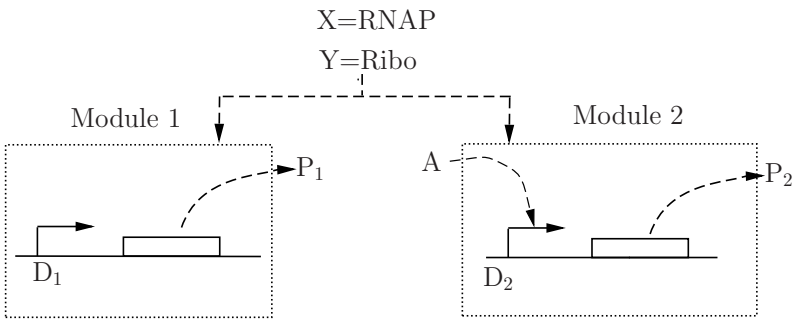


Figure 8.2: Module 1 has a constitutively active promoter that controls the expression of protein P_1, while Module 2 has a promoter activated by activator A, which controls the expression of protein P_2. The two modules do not share any transcription factors, so they are not "connected". Both of them use RNAP (X) and ribosomes (Y) for the transcription and translation processes.

nomenon, we first focus on the usage of RNAP, neglecting the usage of ribosomes. We then provide a computational model that accounts for both RNAP and ribosome utilization and illustrate quantitative simulation results.

Analytical study

To illustrate the essence of the problem, we assume that gene expression is a one-step process, in which the RNA polymerase binds to the promoter region of a gene resulting in a transcriptionally active complex, which, in turn, produces the protein. That is, we will be using the lumped reactions (2.9), in which on the right-hand side of the reaction we have the protein instead of the mRNA.

By virtue of this simplification, we can write the reactions describing Module 1 as:

$$\mathrm{D}_1 + \mathrm{X} \underset{d_1}{\overset{a_1}{\rightleftharpoons}} \mathrm{D}_1{:}\mathrm{X} \xrightarrow{k_1} \mathrm{D}_1 + \mathrm{X} + \mathrm{P}_1, \qquad \mathrm{P}_1 \xrightarrow{\gamma} \emptyset.$$

The reactions describing Module 2 can be written similarly recalling that in the presence of an activator the reactions modify according to (2.18). Taking this into account, the reactions of Module 2 are given by

$$\mathrm{A} + \mathrm{D}_2 \underset{d_0}{\overset{a_0}{\rightleftharpoons}} \mathrm{A}{:}\mathrm{D}_2, \qquad \mathrm{A}{:}\mathrm{D}_2 + \mathrm{X} \underset{d_2}{\overset{a_2}{\rightleftharpoons}} \mathrm{A}{:}\mathrm{D}_2{:}\mathrm{X} \xrightarrow{k_2} \mathrm{A}{:}\mathrm{D}_2 + \mathrm{X} + \mathrm{P}_2, \qquad \mathrm{P}_2 \xrightarrow{\gamma} \emptyset.$$

We let $D_{\mathrm{tot},1}$ and $D_{\mathrm{tot},2}$ denote the total concentration of DNA for Module 1 and Module 2, respectively, and we let $K_\mathrm{d} = d_0/a_0$, $K_1 = d_1/a_1$, and $K_2 = d_2/a_2$. By approximating the complexes concentrations with their quasi-steady state values as illustrated in Section 2.3, we obtain the expressions

$$[\mathrm{D}_1{:}\mathrm{X}] = D_{\mathrm{tot},1}\frac{X/K_1}{1+X/K_1}, \qquad [\mathrm{A}{:}\mathrm{D}_2{:}\mathrm{X}] = D_{\mathrm{tot},2}\frac{(A/K_\mathrm{d})(X/K_2)}{1+(A/K_\mathrm{d})(1+X/K_2)}. \tag{8.1}$$

As a consequence, the differential equation model for the system is given by

$$\frac{dP_1}{dt} = k_1 D_{\text{tot},1} \frac{X/K_1}{1 + X/K_1} - \gamma P_1$$
$$\frac{dP_2}{dt} = k_2 D_{\text{tot},2} \frac{(A/K_\text{d})(X/K_2)}{1 + (A/K_\text{d})(1 + X/K_2)} - \gamma P_2,$$

so that the steady state values of P_1 and P_2 are given by

$$P_1 = \frac{k_1 D_{\text{tot},1}}{\gamma} \frac{X/K_1}{1 + X/K_1}, \qquad P_2 = \frac{k_2 D_{\text{tot},2}}{\gamma} \frac{(A/K_\text{d})(X/K_2)}{1 + (A/K_\text{d})(1 + X/K_2)}.$$

These two values are indirectly coupled through the conservation law of RNAP. Specifically, we let X_{tot} denote the total concentration of RNAP. This value is mainly determined by the cell growth rate and for a given growth rate it is about constant. Then, we have that $X_{\text{tot}} = X + [\text{D}_1\text{:X}] + [\text{A:D}_2\text{:X}]$, which, considering the expressions of the quasi-steady state values of the complexes concentrations in (8.1), leads to

$$X_{\text{tot}} = X + D_{\text{tot},1} \frac{X/K_1}{1 + X/K_1} + D_{\text{tot},2} \frac{(A/K_\text{d})(X/K_2)}{1 + (A/K_\text{d})(1 + X/K_2)}. \tag{8.2}$$

We next study how the steady state value of X is affected by the activator concentration A and how this effect reflects into a dependency of P_1 on A. To perform this study, it is useful to write $\alpha := (A/K_\text{d})$ and note that for α sufficiently small (sufficiently small amounts of activator A), we have that

$$(\alpha(X/K_2))/(1 + \alpha(1 + X/K_2)) \approx \alpha(X/K_2).$$

Also, to simplify the derivations, we assume that the binding of X to D_1 is sufficiently weak, that is, $X \ll K_1$. In light of this, we can re-write the conservation law (8.2) as

$$X_{\text{tot}} = X + D_{\text{tot},1} \frac{X}{K_1} + D_{\text{tot},2} \alpha \frac{X}{K_2}.$$

This equation can be explicitly solved for X to yield

$$X = \frac{X_{\text{tot}}}{1 + (D_{\text{tot},1}/K_1) + \alpha(D_{\text{tot},2}/K_2)}.$$

This expression depends on α, and hence on the activator concentration A. Specifically, as the activator is increased, the value of free X concentration monotonically decreases. As a consequence, the equilibrium value P_1 will also depend on A according to

$$P_1 = \frac{k_1 D_{\text{tot},1}}{\gamma} \frac{X_{\text{tot}}/K_1}{1 + (D_{\text{tot},1}/K_1) + \alpha(D_{\text{tot},2}/K_2)},$$

so that also P_1 monotonically decreases as A is increased. That is, Module 1 responds to changes in the activator of Module 2. From these expressions, we can also deduce that if $D_{\text{tot},1}/K_1 \gg \alpha D_{\text{tot},2}/K_2$, that is, the demand for RNAP in Module 1 is much larger than that of Module 2, then changes in the activator concentration will lead to small changes in the free amount of RNAP and in P_1.

This analysis illustrates that forcing an increase in the expression of any protein causes an overall decrease in available resources, which leads to decrease of expression of other proteins. As a consequence there is a tradeoff between how much protein we can have in a circuit, which is crucial, for example, for the insulation devices designs, and how much the expression of other circuit proteins is reduced. In addition to a design tradeoff, this analysis illustrates that "unconnected" circuits can affect each other because they share common resources. This can, in principle, lead to a dramatic departure of a circuit's behavior from its nominal one.

As an exercise, the reader can verify that similar results would hold in the case in which Module 2 has a repressible promoter instead of one that can be activated (see Exercise 8.1).

The model that we have presented here contains many simplifications. In addition to the mathematical approximations performed and to the fact that it does not account for ribosomes, it does not account for the transcription of endogenous genes. In fact, RNAP is also being used for transcription of chromosomal genes. While the qualitative behavior of the coupling between Module 1 and Module 2 is not going to be affected by including endogenous transcription, the extent of this coupling may be substantially impacted by endogenous transcription. In particular, the quantitative impact of endogenous transcription on this coupling highly depends on the effective demand for RNAP of endogenous genes. This is briefly illustrated in the next section.

Estimates of the effects of adding external plasmids on the availability of RNAP

In the previous section, we illustrated qualitatively the mechanism by which the change in the availability of a shared resouce, due to the addition of synthetic circuits, can cause unexpected crosstalk between unconnected circuits. The extent of this crosstalk depends on the amount by which the shared resource changes. This amount, in turn, depends on the specific values of the dissociation constants, the total resource amounts, and the fraction of resource that is used already by natural circuits.

In *E. coli*, the amount of RNA polymerase and its partitioning mainly depends on the growth rate of the cell [13]: with 0.6 doublings/hour there are only 1500 molecules/ cell, but with 2.5 doublings/hour this number is 11400. The fraction of active RNA polymerase molecules also increases with the growth rate. For illustration purposes, we assume here that the growth rate is the lowest considered in [13].

Therefore, a reasonable estimate is that the total number of RNA polymerase is about 1000. Since the fraction of immature core enzyme at low growth rate is only a few percent [14], we assume that the total number of functional RNA polymerase is about 1000 per cell, that is, we set $X_{\text{tot}} = 1000\text{nM}$ (recalling that 1 molecule corresponds to a concentration of about 1.6 nM). Based on the data presented in [14], a reasonable partitioning of RNA polymerase is the following:

active core enzyme: 15% (150 molecules/cell or $X_{\text{a}} = 150\text{nM}$),

promoter-bound holoenzyme: 15% (150 molecules/cell or $X_{\text{p}} = 150\text{nM}$),

free holoenzyme: 5% (50 molecules/cell or $X_{\text{f}} = 50\text{nM}$),

inactive DNA-bound core: 65% (650 molecules/cell $X_{\text{i}} = 650\text{nM}$).

There are about 1000 genes expressed in exponential growth phase [44], hence we approximate the number of binding sites for X to 1000, or $D_{\text{tot}} = 1000\text{nM}$, and we assume that all the 150 promoter-bound holoenzymes are bound to these promoters. The binding reaction is of the form

$$\text{D} + \text{X}_{\text{f}} \underset{d}{\overset{a}{\rightleftharpoons}} \text{D:X},$$

in which D is the DNA promoter in total concentration D_{tot}. Consequently, we have $D_{\text{tot}} = D + [D : X]$. At the equilibrium, we have $[D : X] = X_{\text{p}} = 150\text{nM}$ and $D = D_{\text{tot}} - [D : X] = D_{\text{tot}} - X_{\text{p}} = 850\text{nM}$. With dissociation constant $K_{\text{d}} = \frac{d}{a}$ the equilibrium is given by $0 = X_{\text{f}}D - K_{\text{dD:X}}$, hence we have that

$$K_{\text{d}} = \frac{D}{[D : X]} X_{\text{f}} \approx 300\text{nM},$$

which is interpreted as an "effective" dissociation constant. This is in the range $1 - 1000\text{nM}$ suggested by [36] for specific binding of RNA polymerase to DNA. Therefore, we are going to model the binding of RNA polymerase to the promoters of the chromosome of *E. coli* in exponential phase as one promoter with concentration D_{tot} and effective dissociation constant K_{d}.

Furthermore, we have to take into account the rather significant amount of RNA polymerase bound to the DNA other than at the promoter region ($X_{\text{a}} + X_{\text{i}} = 800\text{nM}$). To do so, we assume that the fraction $m = X_{\text{a}} + X_{\text{i}} + X_{\text{p}}/X_{\text{p}}$ is approximately constant at the equilibrium.

Now, we can consider the addition of synthetic plasmids. Specifically, we consider the plasmid pSB1AK3 (copy number $100 - 300$) with one copy of a gene under the control of a constitutive promoter. The binding of RNA polymerase to the constitutive promoter is modeled by

$$\text{D}' + \text{X}_{\text{f}} \underset{d'}{\overset{a'}{\rightleftharpoons}} \text{D}':\text{X}$$

where D' is the RNA polymerase-free promoter and D′:X is the RNA polymerase: promoter complex. Consequently, we have $D'_{\text{tot}} = D' + [D' : X]$. The dissociation constant is given by $K'_{\text{d}} = \frac{d'}{a'}$. The total concentration of promoters D'_{tot} can be determined by considering the copy number of the plasmid, which is $100 - 300$ plasmids/cell, so that we set $D'_{\text{tot}} \approx 200\text{nM}$. At the equilibrium, we have

$$[D' : X] = D'_{\text{tot}} \frac{X_{\text{f}}}{K'_{\text{d}} + X_{\text{f}}}.$$

We also have

$$[D : X] = D_{\text{tot}} \frac{X_{\text{f}}}{K_{\text{d}} + X_{\text{f}}}.$$

The conservation law for RNA polymerase must be now considered in order to determine the equilibrium concentrations:

$$X_{\text{f}} + m\,[D : X] + [D' : X] = X_{\text{tot}}. \tag{8.3}$$

Here, we did not account for RNA polymerase molecules paused, queuing and actively transcribing on the plasmid, moreover, we also neglected the resistance genes on the pSB1AK3 plasmid. Hence, we are underestimating the effect of load presented by the plasmid.

Solving equation (8.3) for the free RNA polymerase amount X_{f} gives the following results. These results depend on the ratio between the effective dissociation constant K_{d} and the dissociation constant K'_{d} of RNA polymerase from the plasmid promoter:

> $K'_{\text{d}} = 0.1K_{\text{d}}$ (RNA polymerase binds better to the plasmid promoter) results in $X_{\text{f}} = 21\text{nM}$, $[D : X] = 69\text{nM}$ and $[D' : X] = 85\text{nM}$. Hence, the concentration of free RNA polymerase decreases by about 60%;
>
> $K'_{\text{d}} = K_{\text{d}}$ (binding is the same) results in $X_{\text{f}} = 41\text{nM}$, $[D : X] = 126\text{nM}$ and $[D' : X] = 25\text{nM}$. Hence, the concentration of free RNA polymerase decreases by about 20%;
>
> $K'_{\text{d}} = 10K_{\text{d}}$ (RNA polymerase binds better to the chromosome) results in $X_{\text{f}} = 49\text{nM}$, $[D : X] = 147\text{nM}$ and $[D' : X] = 3\text{nM}$. Hence, the concentration of free RNA polymerase decreases by about 2%.

We conclude that if the promoter on the synthetic plasmids has a dissociation constant for RNA polymerase that is in the range of the effective one calculated above, the perturbation on the available free RNA polymerase is abut 20%. This perturbation, even if fairly small, may in practice result into large effects on the protein concentration. This is because it may cause a large perturbation in the concentration of free ribosomes. In fact, one added copy of an exogenous plasmid will lead to transcription of several mRNA molecules, which will demand ribosomes

for translation. Hence, a small increase in the demand for RNAP may lead to a dramatically larger increase in the demand for ribosomes. This is illustrated in the next section through a computational model including ribosome sharing.

Computational model and numerical study

In this section, we introduce a model of the system in Figure 8.2, in which we consider both the RNA polymerase and the ribosome usage. We let the concentration or RNA polymerase be denoted by X and the concentration of ribosomes be denoted by Y. We let m_1 and P_1 denote the concentrations of the mRNA and protein in Module 1 and let m_2 and P_2 denote the concentrations of the mRNA and protein in Module 2. The reactions of the transcription process in Module 1 are given by (see Section 2.2):

$$\mathrm{X} + \mathrm{D}_1 \underset{d_1}{\overset{a_1}{\rightleftharpoons}} \mathrm{X{:}D}_1, \quad \mathrm{X{:}D}_1 \xrightarrow{k_1} \mathrm{m}_1 + \mathrm{X} + \mathrm{D}_1, \quad \mathrm{m}_1 \xrightarrow{\delta} \emptyset,$$

while the translation reactions are given by

$$\mathrm{Y} + \mathrm{m}_1 \underset{d_1'}{\overset{a_1'}{\rightleftharpoons}} \mathrm{Y{:}m}_1, \quad \mathrm{Y{:}m}_1 \xrightarrow{k_1'} \mathrm{P}_1 + \mathrm{Y} + \mathrm{m}_1, \quad \mathrm{P}_1 \xrightarrow{\gamma} \emptyset, \quad \mathrm{Y{:}m}_1 \xrightarrow{\delta} \mathrm{Y}.$$

The resulting system of differential equations is given by

$$\begin{aligned}
\frac{d}{dt}[\mathrm{X{:}D}_1] &= a_1 X D_1 - (d_1 + k_1)[\mathrm{X{:}D}_1] \\
\frac{d\, m_1}{dt} &= k_1 [\mathrm{X{:}D}_1] - a_1' Y m_1 + d_1' [\mathrm{Y{:}m}_1] - \delta m_1 + k_1' [\mathrm{Y{:}m}_1] \\
\frac{d}{dt}[\mathrm{Y{:}m}_1] &= a_1' Y m_1 - (d_1' + k_1;)[\mathrm{Y{:}m}_1] \\
\frac{d\, P_1}{dt} &= k_1' [\mathrm{Y{:}m}_1] - \gamma P_1,
\end{aligned} \tag{8.4}$$

in which $D_1 = D_{\mathrm{tot},1} - [\mathrm{X{:}D}_1]$ from the conservation law of DNA in Module 1.

The reactions of the transcription process in Module 2 are given by (see Section 2.3)

$$\mathrm{A} + \mathrm{D}_2 \underset{d_0}{\overset{a_0}{\rightleftharpoons}} \mathrm{A{:}D}_2, \qquad \mathrm{A{:}D}_2 + \mathrm{X} \underset{d_2}{\overset{a_2}{\rightleftharpoons}} \mathrm{A{:}D}_2\mathrm{{:}X} \xrightarrow{k_2} \mathrm{A{:}D}_2 + \mathrm{X} + \mathrm{m}_2, \qquad \mathrm{m}_2 \xrightarrow{\delta} \emptyset,$$

while the translation reactions are given by

$$\mathrm{Y} + \mathrm{m}_2 \underset{d_2'}{\overset{a_2'}{\rightleftharpoons}} \mathrm{Y{:}m}_2, \quad \mathrm{Y{:}m}_2 \xrightarrow{k_2'} \mathrm{P}_2 + \mathrm{Y} + \mathrm{m}_2, \quad \mathrm{Y{:}m}_2 \xrightarrow{\delta} \mathrm{Y} \quad \mathrm{P}_2 \xrightarrow{\gamma} \emptyset.$$

The resulting system of differential equations is given by

$$\begin{aligned}
\frac{d}{dt}[\mathrm{A{:}D_2}] &= a_0\, A\, D_2 - d_0\,[\mathrm{A{:}D_2}] - a_2\, X\,[\mathrm{A{:}D_2}] + (d_2 + k_2)[\mathrm{A{:}D_2{:}X}] \\
\frac{d}{dt}[\mathrm{A{:}D_2{:}X}] &= a_2\, X\,[\mathrm{A{:}D_2}] - (d_2 + k_2)\,[\mathrm{A{:}D_2{:}X}] \\
\frac{d\, m_2}{dt} &= k_2\,[\mathrm{A{:}D_2{:}X}] - a_2'\, Y\, m_2 + d_2'\,[\mathrm{Y{:}m_2}] - \delta\, m_2 + k_2'\,[\mathrm{Y{:}m_2}] \qquad (8.5) \\
\frac{d}{dt}[\mathrm{Y{:}m_2}] &= a_2'\, Y\, m_2 - (d_2' + k_2')\,[\mathrm{Y{:}m_2}] - \delta[\mathrm{Y{:}m_2}] \\
\frac{d\, P_2}{dt} &= k_2'\,[\mathrm{Y{:}m_2}] - \gamma\, P_2,
\end{aligned}$$

in which, we have that $D_2 = D_{\mathrm{tot},2} - [\mathrm{A{:}D_2}] - [\mathrm{A{:}D_2{:}X}]$ by the conservation law of DNA in Module 2.

The two modules are coupled by the conservation laws for RNAP and ribosomes given by

$$X_{\mathrm{tot}} = X + [\mathrm{X{:}D_1}] + [\mathrm{A{:}D_2{:}X}], \quad Y_{\mathrm{tot}} = Y + [\mathrm{Y{:}m_1}] + [\mathrm{Y{:}m_2}],$$

which are employed in systems (8.4)-(8.5) by writing

$$X = X_{\mathrm{tot}} - [\mathrm{X{:}D_1}] - [\mathrm{A{:}D_2{:}X}], \quad Y = Y_{\mathrm{tot}} - [\mathrm{Y{:}m_1}] - [\mathrm{Y{:}m_2}].$$

The results are shown in Figure 8.3a-8.3d. In the simulations, we have chosen $X_{\mathrm{tot}} = 1\mu$M to account for the fact that the total amount of RNAP in wild type cells at fast division rate i is given by about 10μM of which only 1μM is free, while the rest is bound to the endogenous DNA. Since in the simulations we did not account for endogenous DNA, we assumed that only 1μM is available in total to the two exogenous modules. A similar reasoning was employed to set $Y_{\mathrm{tot}} = 10\mu$M. In exponential growth, we have about 34μM of total ribosomes concentration, but only about 30% of this is free, resulting in about 10μM concentration of ribosomes available to the exogenous modules.

Figure 8.3a illustrates that as the activator concentration A increases, there is no substantial perturbation on the free amount of RNAP. However, because the resulting perturbation on the free amount of ribosomes (8.3a) is significant, the resulting decrease of P_1 is substantial.

8.2 Stochastic Effects: Design Tradeoffs in Systems with Large Gains

[1]As we have seen in Chapter 7, a biomolecular system can be rendered insensitive to retroactivity by implementing a large input amplification gain in a negative

[1]This section is extracted from Jayanthi and Del Vecchio CDC 2009.

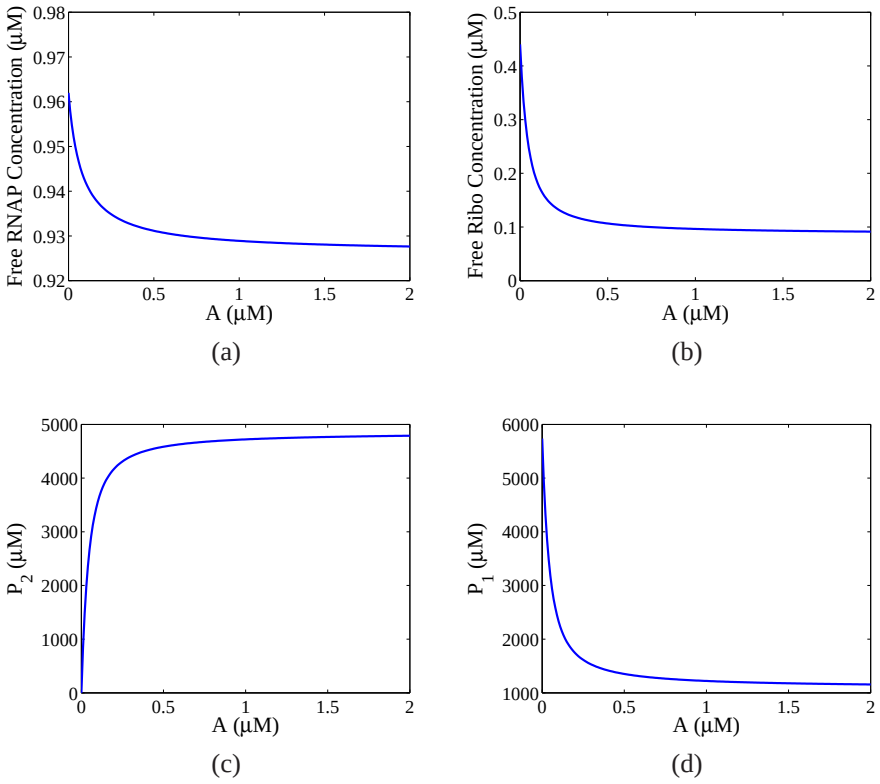


Figure 8.3: Simulation results for the ordinary differential equation model (8.5)-(8.4). For this model, the parameter values were taken from `http://bionumbers.hms.harvard.edu/` as follows. For the concentrations, we have $X_{\text{tot}} = 1\mu\text{M}$, $Y_{\text{tot}} = 10\mu\text{M}$, $D_{\text{tot},1} = D_{\text{tot},2} = 0.2\mu\text{M}$. The values of the association and dissociation rate constants were chosen such that the corresponding dissociation constants were in the range of specific binding dissociation constants. Specifically, we have $a_0 = 10\mu\text{M}^{-1}\text{min}^{-1}$, $d_0 = 1\text{min}^{-1}$, $a_2 = 10\mu\text{M}^{-1}\text{min}^{-1}$, $d_2 = 1\text{min}^{-1}$, $a_2' = 100\mu\text{M}^{-1}\text{min}^{-1}$, $d_2' = 1\text{min}^{-1}$, $a_1 = 10\mu\text{M}^{-1}\text{min}^{-1}$, $d_1 = 1\text{min}^{-1}$, $a_1' = 10\mu\text{M}^{-1}\text{min}^{-1}$, and $d_1' = 1\text{min}^{-1}$. The transcription and translation rate constants were calculated based on the elongation speeds, the average length of a gene, and the average number of RNAP per gene and of ribosomes per transcript. The resulting values chosen are given by $k_1 = k_2 = 40\text{min}^{-1}$ and $k_1' = k_2' = 6\text{min}^{-1}$. Finally, the decay rates are given by $\gamma = 0.01\text{min}^{-1}$ corresponding to a protein half life of about 70 minutes and $\delta = 0.1\text{min}^{-1}$ corresponding to a mRNA half life of about 7 minutes.

feedback loop. However, relying on a large negative feedback, this type of design may have undesired effects in the presence of noise as seen in a different context in Section 6.2. Also, it is not clear so far what the effect of retroactivity is on the noise content of the upstream system. Here, we employ the Langevin equation introduced in Chapter 4 to address these questions.

Consider a transcription system that takes a transcription factor U as input and produces a transcription factor Z as output. The transcription rate of the gene z, which expresses the protein Z, is given by a time varying function $Gk(t)$ that depends on the transcription factor U. This dependency is not modeled, since it is not central to our discussion. The parameter G models the input amplification gain. The degradation rate of protein Z is also assumed to be tunable and thus identified by $G\gamma$. The variable gain parameter G will be adjusted to improve the retroactivity attenuation.

The transcription factor Z is also an input to the downstream load through the reversible binding of Z to promoter sites p. Neglecting the Z messenger RNA dynamics, the system can be modeled by the chemical equations

$$0 \underset{G\gamma}{\overset{G\,k(t)}{\rightleftharpoons}} \mathrm{Z}, \qquad \mathrm{Z} + \mathrm{p} \underset{k_{\mathrm{off}}}{\overset{k_{\mathrm{on}}}{\rightleftharpoons}} \mathrm{C}.$$

We assume that $k(t)$ and γ are of the same order and denote $K_{\mathrm{d}} = k_{\mathrm{off}}/k_{\mathrm{on}}$. We also assume that the production and decay processes are slower than binding and unbinding reactions, that is, $k_{\mathrm{off}} \gg G\gamma$, $k_{\mathrm{on}} \gg G\gamma$ as performed before. Let the total concentration of promoter be p_{tot}. The deterministic ordinary differential equation model is given by

$$\begin{aligned} \frac{dZ}{dt} &= Gk(t) - G\gamma Z + k_{\mathrm{off}} Z - k_{\mathrm{on}}(p_{\mathrm{tot}} - C)Z, \\ \frac{dC}{dt} &= -k_{\mathrm{off}} C + k_{\mathrm{on}}(p_{\mathrm{tot}} - C)Z. \end{aligned} \tag{8.6}$$

To identify by what amounts G should be increased to compensate the retroactivity effect, we perform a linearized analysis of (8.6) about $k(t) = \bar{k}$, and the corresponding equilibrium $\bar{Z} = \bar{k}/\gamma$ and $\bar{C} = \bar{Z}p_{\mathrm{tot}}/(\bar{Z} + K_{\mathrm{d}})$. By performing the linearized analysis as in Section 7.3, letting $z = Z - \bar{Z}$ and $\tilde{k} = k - \bar{k}$, we obtain

$$\frac{dz}{dt} = \frac{G}{1 + R_l}(\tilde{k}(t) - \gamma z), \qquad R_l = \frac{K_{\mathrm{d}} p_{\mathrm{tot}}}{(\bar{k}/\gamma + K_{\mathrm{d}})^2}. \tag{8.7}$$

Thus, we should choose $G \approx 1 + R_l$ to compensate for retroactivity from the load. In real systems, however, there are practical limitations on how much the gain can be increased so that retroactivity may not be completely rejected.

We have shown that increasing the gain G is beneficial for rejecting retroactivity to the upstream component. However, as shown in Figure 8.4, increasing the gain G impacts the frequency content of the noise in a single realization. For low values of G, the error signal between a realization and the mean is of lower frequency when compared to a higher gain.

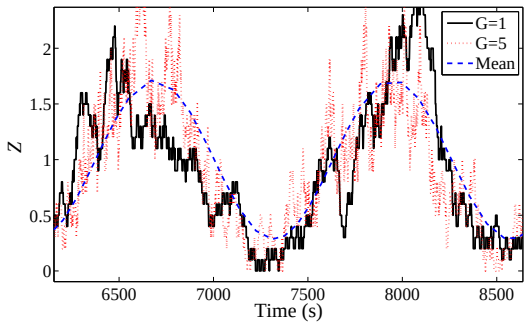


Figure 8.4: Increasing the value of G produces a disturbance signal of higher frequency. Two realizations are shown with different values for G without load. The parameters used in the simulations are $\gamma = 0.01\text{s}^{-1}$, $K_\text{d} = 20\text{nM}$, $k_\text{off} = 50\text{s}^{-1}$, $\omega = 0.005\text{rad/s}$ and $\Omega = 10\text{nM}^{-1}$. The input signal used is $k(t) = \gamma(1 + 0.8\sin(\omega t))\text{s}^{-1}$. The mean of the signal is given as reference.

To study this problem, we employ the Langevin equation (Section 4.1). For our system, we obtain

$$\begin{aligned}\frac{dZ}{dt} =& Gk(t) - G\gamma Z - k_\text{on}(p_\text{tot} - C)Z + k_\text{off}C + \sqrt{Gk(t)}\,N_1(t) - \sqrt{G\gamma Z}\,N_2(t) \\ &- \sqrt{k_\text{on}(p_\text{tot} - C)Z}\,N_3(t) + \sqrt{k_\text{off}C}\,N_4(t), \\ \frac{dC}{dt} =& k_\text{on}(p_\text{tot} - C)Z - k_\text{off}C + \sqrt{k_\text{on}(p_\text{tot} - C)Z}\,N_3(t) - \sqrt{k_\text{off}C}\,N_4(t).\end{aligned} \tag{8.8}$$

The above system can be viewed as a non-linear system with five inputs, $k(t)$ and $N_i(t)$ for $i = 1, 2, 3, 4$. Let $k(t) = \bar{k}$, $N_1(t) = N_2(t) = N_3(t) = N_4(t) = 0$ be constant inputs and let $\bar{Z}$ and $\bar{C}$ be the corresponding equilibrium points. Then for small amplitude signals $\tilde{k}(t)$ the linearization of the system (8.8) leads, with abuse of notation, to

$$\begin{aligned}\frac{dZ}{dt} =& G\tilde{k}(t) - G\gamma Z - k_\text{on}(p_\text{tot} - \bar{C})Z + k_\text{on}\bar{Z}C + k_\text{off}C \\ &+ \sqrt{G\bar{k}}\,N_1(t) - \sqrt{\gamma\bar{Z}}\,N_2(t) - \sqrt{k_\text{off}\bar{C}}\,N_3(t) + \sqrt{k_\text{on}(p_\text{tot} - \bar{C})\bar{Z}}\,N_4(t) \\ \frac{dC}{dt} =& k_\text{on}(p_\text{tot} - \bar{C})Z - k_\text{on}\bar{Z}C - k_\text{off}C + \sqrt{k_\text{off}\bar{C}}\,N_3(t) - \sqrt{k_\text{on}(p_\text{tot} - \bar{C})\bar{Z}}\,N_4(t).\end{aligned} \tag{8.9}$$

We can further simplify the above expressions by noting that $\gamma\bar{Z} = G\bar{k}$ and $k_\text{on}(p_\text{tot} - \bar{C})\bar{Z} = k_\text{off}\bar{C}$. Also, since N_j are independent identical Gaussian white noises, we can write $N_1(t) - N_2(t) = \sqrt{2}\Gamma_1(t)$ and $N_3(t) - N_4(t) = \sqrt{2}\Gamma_2(t)$, in which $\Gamma_1(t)$ and $\Gamma_2(t)$ are independent Gaussian white noises identical to $N_j(t)$. This simplification

leads to the system

$$\frac{dZ}{dt} = G\tilde{k}(t) - G\gamma Z - k_{\text{on}}(p_{\text{tot}} - \bar{C})Z + k_{\text{on}}\bar{Z}C + k_{\text{off}}C + \sqrt{2G\bar{k}}\Gamma_1(t) - \sqrt{2k_{\text{off}}\bar{C}}\Gamma_2(t),$$
$$\frac{dC}{dt} = k_{\text{on}}(p_{\text{tot}} - \bar{C})Z - k_{\text{on}}\bar{Z}C - k_{\text{off}}C + \sqrt{2k_{\text{off}}\bar{C}}\Gamma_2(t). \tag{8.10}$$

This is a system with three inputs: the deterministic input $\tilde{k}(t)$ and two independent white noise sources $\Gamma_1(t)$ and $\Gamma_2(t)$. One can interpret Γ_1 as the source of the fluctuations caused by the production and degradation reactions while Γ_2 is the source of fluctuations caused by binding and unbinding reactions. Since the system is linear, we can analyze the different contributions of each noise source separately and independent from the signal $\tilde{k}(t)$.

The transfer function from Γ_1 to Z is (after setting $\gamma/k_{\text{off}} = \epsilon = 0$)

$$H_{Z\Gamma_1}(s) = \frac{\sqrt{2Gk}}{s(1+R_l)+G\gamma}. \tag{8.11}$$

The zero frequency gain of this transfer function is equal to $H_{Z\Gamma_1}(0) = \sqrt{2k}/\sqrt{G}\gamma$. Thus, as G increases, the zero frequency gain decreases. But for large enough frequencies ω, $j\omega(1+R_l)+G\gamma \approx j\omega(1+R_l)$, and the amplitude $|H_{Z\Gamma_1}(j\omega)| \approx \sqrt{2kG}/\omega(1+R_l)$ becomes a monotone function of G. This effect is illustrated in Figure 8.5. The frequency at which the amplitude of $|H_{Z\Gamma_1}(j\omega)|$ computed with $G = 1$ intersects the amplitude $|H_{Z\Gamma_1}(j\omega)|$ computed with $G > 1$ is given by the expression

$$\omega_e = \frac{\gamma\sqrt{G}}{(1+R_l)}.$$

Thus, when increasing the gain from 1 to $G > 1$, we reduce the noise at frequencies lower than ω_e but we increase it at frequencies larger than ω_e. Note , in particular, that there is an increase of the amplitude at the frequency of interest $\omega = 0.01$.

While retroactivity contributes to filtering noise in the upstream system as it decreases the bandwidth of the noise transfer function, high gains contribute to increasing noise at frequencies higher than ω_e. In particular, when increasing the gain from 1 to $G > 1$ we reduce the noise in the frequency ranges below $\omega_e = \gamma\sqrt{G}/(R_l+1)$, but the noise at frequencies above ω_e increases. If we were able to indefinitely increase G, we could send G to infinity attenuating the deterministic effects of retroactivity while amplifying noise only at very high, hence not relevant, frequencies.

In practice, however, the value of G is limited. For example, in the insulation device based on phosphorylation, G is limited by the amounts of substrate and phosphatase that we can have in the system. Hence, a design tradeoff needs to be considered when designing insulation devices: placing the largest possible G attenuates retroactivity but it may increase noise in a possibly relevant frequency range.

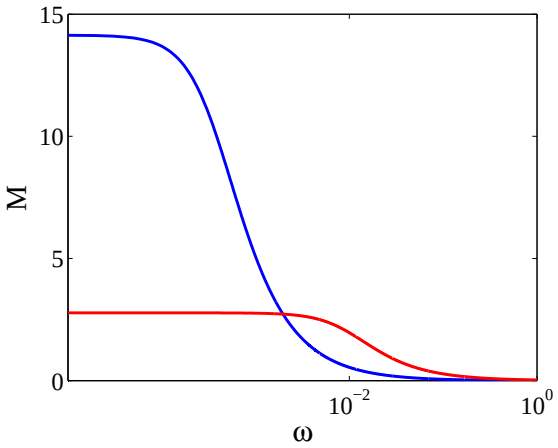


Figure 8.5: Magnitude M of the transfer functions $H_{Z\Gamma_1}(s)$ as a function of the input frequency ω. The parameters used in this plot are $\gamma = 0.01\text{s}^{-1}$, $K_\text{d} = 1\text{nM}$, $k_\text{off} = 50\text{s}^{-1}$, $\omega = 0.005\text{rad/s}$, $p_\text{tot} = 100\text{nM}$. When G increases from 1 to $1 + R_l = 25$, the contribution from Γ_1 decreases at low frequency but it spreads to a higher range of the frequency.

Exercises

8.1 In the case of a repressor, a similar derivation to what was performed in the text can be carried if R were a repressor of the transcription of Module 2. Using a one-step reaction model for gene expression, write down the reaction equations for this case and the reaction rate equations describing the rate of change of P_1 and P_2. Then, determine how the free concentration of RNAP is affected by changes in R and how P_1 is affected by changes in R.

8.2 Consider again the case of a repressor as considered in the previous example. Now, consider a two-step reaction model for transcription and build a simulation model with parameter values as indicated in the text and determine the extent of coupling between Module 1 and Module 2 when the repressor is increased.

Appendix A
A Primer on Control Theory

This appendix provides a brief primer on some of the key topics in control theory that are used in the text. The material here is drawn from *Feedback Systems* by Åström and Murray.

A.1 System Modeling

A model is a precise representation of a system's dynamics used to answer questions via analysis and simulation. The model we choose depends on the questions we wish to answer, and so there may be multiple models for a single physical system, with different levels of fidelity depending on the phenomena of interest. In this chapter we provide an introduction to the concept of modeling, and provide some basic material on two specific methods that are commonly used in feedback and control systems: differential equations and difference equations.

1. A *model* is a mathematical representation of a system that can be used to answer question about that system. The choice of the model depends on the questions one wants to ask. Models for control systems are typically input/output models and combine techniques from mechanics and electrical engineering.

2. The *state* of a system is a collection of variables that summarize the past history of the system for the purpose of predicting the future. A *state space model* is one that describe how the state of a system evolves over time.

3. We can model the evolution of the state using a *ordinary differential equations* of the form

$$\begin{aligned} \dot{x} &= f(x,u) & \dot{x} &= Ax + Bu \\ y &= h(x,u) & y &= Cx + Du \end{aligned} \tag{A.1}$$

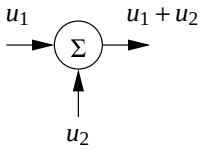
   where x represents the state of the system, $\dot{x}$ is the time derivative of the state, u are the external inputs and y are the measured outputs. For the linear form, A, B, C and D are matrices of the appropriate dimension and the model is *linear time invariant* (LTI).

4. Three common questions that can be answered using state space models are (1) how the system state evolves from a given initial condition, (2) the stabil-

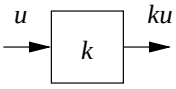
ity of an equilibrium point from nearby initial conditions and (3) the steady state response of the system to sinusoidal forcing at different frequencies.

5. Models can be constructed from experiments by measuring the response of a system and determining the parameters in the model that correspond to features in the response. Examples include measuring the period of oscillation, the rate of damping and the steady state amplitude of the response of a system to a step input.

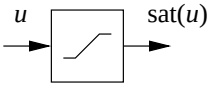
6. Schematic and block diagrams are common tools for modeling large, complex systems. The following symbols are some of the ones commonly used for modeling control systems:



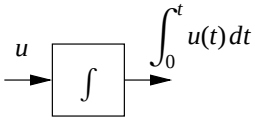
Summing junction



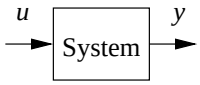
Gain block



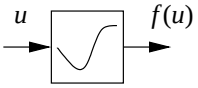
Saturation



Integrator



Input/output system



Nonlinear map

Computer packages such as LabView, MATLAB/SIMULINK and Modelica can be used to construct models for complex, multi-component systems.

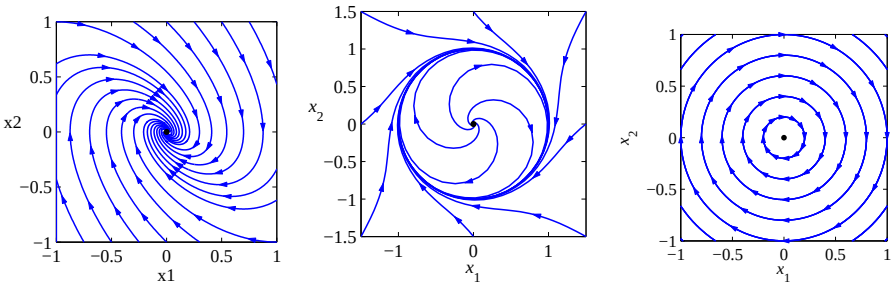
A.2 Dynamic Behavior

In this chapter we give a broad discussion of the behavior of dynamical systems, focused on systems modeled by nonlinear differential equations. This allows us to discuss equilibrium points, stability, limit cycles and other key concepts of dynamical systems. We also introduce some methods for analyzing global behavior of solutions.

1. We say that $x(t)$ is a solution of a differential equation on the time interval t_0 to t_f with initial value x_0 if it satisfies

$$x(t_0) = x_0 \quad \text{and} \quad \dot{x}(t) = F(x(t)) \quad \text{for all} \quad t_0 \le t \le t_f. \tag{A.2}$$

We will usually assume $t_0 = 0$. For most differential equations we will encounter, there is a unique solution for a given initial condition. Numerical



(a) Asymptotically stable (b) Unstable (c) Stable

Figure A.1: Basic features of dynamical systems. (a) An asymptotically stable equilibrium point at $x = (0,0)$. (b) A limit cycle of radius one, with an unstable equilibrium point at $x = (0,0)$. (c) A stable equlibirum point at $x = (0,0)$ (nearby initial conditions stay nearby).]

tools such as MATLAB and Mathematica can be used to obtain numerical solutions for $x(t)$ given the function $F(x)$.

2. An *equilibrium point* for a dynamical system represents a point x_e such that if $x(0) = x_e$ then $x(t) = x_e$ for all t. Equilibrium points represent stationary conditions for the dynamics of a system. A *limit cycle* for a dynamical system is a solution $x(t)$ which is periodic with some period T, so that $x(t+T) = x(t)$ for all t.

3. An equilibrium point is (locally) *stable* if initial conditions that start near an equilibrium point stay near that equilibrium point. A equilibrium point is (locally) *asymptotically stable* if it is stable and, in addition, the state of the system converges to the equilibrium point as time increases. An equilibrium point is *unstable* if it is not stable. Similar definitions can be used to define the stability of a limit cycle.

4. Phase portraits provide a convenient way to understand the behavior of 2-dimensional dynamical systems. A phase portrait is a graphical representation of the dynamics obtained by plotting the state $x(t) = (x_1(t), x_2(t))$ in the plane. This portrait is often augmented by plotting an arrow in the plane corresponding to $F(x)$, which shows the rate of change of the state. Figure A.1 illustrates some of the basic features of a dynamical systems.

5. A linear system

$$\frac{dx}{dt} = Ax \tag{A.3}$$

is asymptotically stable if and only if all eigenvalues of A all have strictly negative real part and is unstable if any eigenvalue of A has strictly positive real part. A nonlinear system can be approximated by a linear system around

an equilibrium point by using the relationship

$$\dot{x} = F(x_e) + \left.\frac{\partial F}{\partial x}\right|_{x_e} (x - x_e) + \text{higher order terms in } (x - x_e). \tag{A.4}$$

Since $F(x_e) = 0$, we can approximate the system by choosing a new state variable $z = x - x_e$ and writing the dynamics as $\dot{z} = Az$. The stability of the nonlinear system can be determined in a local neighborhood of the equilibrium point through its linearization.

6. A *Lyapunov function* is an energy-like function $V : R^n \to R$ that can be used to reason about the stability of an equilibrium point. We define the derivative of V along the trajectory of the system as

$$\dot{V}(x) = \frac{\partial V}{\partial x}\dot{x} = \frac{\partial V}{\partial x}F(x) \tag{A.5}$$

Assuming $x_e = 0$ and $V(0) = 0$, the following conditions hold:

Condition on V	Condition on $\dot{V}$	Stability
$V(x) > 0, x \neq 0$	$\dot{V}(x) \leq 0$ for all x	x_e stable
$V(x) > 0, x \neq 0$	$\dot{V}(x) < 0, x \neq 0$	x_e asymptotically stable

Stability of limit cycles can also be studied using Lyapunov functions.

7. The *global behavior* of a nonlinear system refers to dynamics of the system far away from equilibrium points. The *region of attraction* of an asymptotically stable equilirium point refers to the set of all initial conditions that converge to that equilibrium point. An equilibrium point is said to be *globally asymptotically stable* if all initial conditions converge to that equilibrium point. Global stability can be checked by finding a Lyapunov function that is globally positive definition with time derivative globally negative definite.

A.3 Linear Systems

Previous chapters have focused on the dynamics of a system with relatively little attention to the inputs and outputs. This chapter gives an introduction to input/output behavior for linear systems and shows how a nonlinear system can be approximated near an equilibrium point by a linear model.

1. A *linear system* is one in which the output is jointly linear in the intitial condition for the system and the input to the system. In particular, a linear system has the property that if we apply an input $u(t) = \alpha u_1(t) + \beta u_2(t)$ with zero initial condition, the corresponding output will be $y(t) = \alpha y_1(t) + \beta y_2(t)$, where y_i is the output associated with the input u_i. This propery is called linear *superposition*.

2. A differential equation of the form

$$\begin{aligned} \dot{x} &= Ax + Bu \qquad x \in R^n, u \in R \\ y &= Cx + Du \qquad y \in R \end{aligned} \tag{A.6}$$

is a *single-input, single-output* (SISO) *linear differential equation*. Its solution can be written in terms of the *matrix exponential*

$$e^{At} = I + At + \frac{1}{2}A^2t^2 + \frac{1}{3!}A^3t^3 + \cdots = \sum_{k=0}^{\infty} \frac{1}{k!}A^k t^k. \tag{A.7}$$

The solution to the differential equation is given by the *convolution equation*

$$y(t) = Ce^{At}x(0) + \int_0^t Ce^{A(t-\tau)}Bu(\tau)d\tau + Du(t). \tag{A.8}$$

3. A linear system

$$\dot{x} = Ax \tag{A.9}$$

is *asymptotically stable* if and only if all eigenvalues of A all have strictly negative real part and is unstable if any eigenvalue of A has strictly positive real part. For systems with eigenvalues having zero real-part, stability is determined by using the Jordan normal form associated with the matrix. A system with eigenvalues that have no strictly positive real part is stable if and only if the Jordan block corresponding to each eigenvalue with zero part is a scalar (1x1) block.

4. The input/output response of a (stable) linear system contains a transient region portion, which eventually decays to zero, and a steady state portion, which persists over time. Two special responses are the *step response*, which is the output corresponding to an step input applied at $t = 0$ and the *frequency response*, which is the response of the system to a sinusoidal input at a given frequency.

5. The step response is characterized by the following parameters:

   - The *steady state value*, y_{ss}, of a step response is the final level of the output, assuming it converges.
   - The *rise time*, T_r, is the amount of time required for the signal to go from 10value.
   - The *overshoot*, M_p, is the percentage of the infal value by which the signal initially rises above the final value.
   - The *settling time*, T_s, is the amount of time required for the signal to stay within 5times.

6. The frequency response is given by

$$y(t) = \underbrace{Ce^{At}\left(x(0) - (sI - A)^{-1}B\right)}_{\text{transient}} + \underbrace{\left(D + C(sI - A)^{-1}B\right)e^{st}}_{\text{steady state}}, \tag{A.10}$$

where $\cos\omega t = \frac{1}{2}\left(e^{j\omega t} + e^{-j\omega t}\right)$ and $s = j\omega$. The gain and phase of the frequency response are given by

$$\text{gain}(\omega) = \frac{A_y}{A_u} = M \qquad \text{phase}(\omega) = \phi - \psi = \theta. \tag{A.11}$$

7. A nonlinear system of the form

$$\begin{aligned} \dot{x} &= f(x,u) \qquad x \in R^n, u \in R \\ y &= h(x,u) \qquad y \in R \end{aligned} \tag{A.12}$$

is a single-input, single-output (SISO) nonlinear system. It can be linearized about an equibrium point $x = x_e$, $u = u_e$, $y = y_e$ by defining new variables

$$z = x - x_e \qquad v = u - u_e \qquad w = y - h(x_e, u_e). \tag{A.13}$$

The dynamics of the system near the equilibrium point can then be approximated by the linear system

$$\begin{aligned} \dot{x} &= Ax + Bu \\ y &= Cx + Du \end{aligned} \tag{A.14}$$

where

$$\begin{aligned} A &= \left.\frac{\partial f(x,u)}{\partial x}\right|_{x_e,u_e} \qquad B = \left.\frac{\partial f(x,u)}{\partial u}\right|_{x_e,u_e} \\ C &= \left.\frac{\partial h(x,u)}{\partial x}\right|_{x_e,u_e} \qquad D = \left.\frac{\partial y(x,u)}{\partial u}\right|_{x_e,u_e} \end{aligned} \tag{A.15}$$

The equilibrium point for a nonlinear system is locally asymptotically stable if the real part of the eigenvalues of the linearization about that equilibrium point have strictly negative real part.

A.4 Reachability and observability

The concept of reachability is introduced and used to investigate how to "design" the dynamics of a system through placement of its eigenvalues. In particular, it will be shown that under certain conditions it is possible to assign the system eigenvalues to arbitrary values by appropriate feedback of the system state. We introduce the concept of observability and show that if a system is observable, it is possible to recover the state from measurements of the inputs and outputs to the system.

1. A linear system with dynamics

$$\begin{aligned} \dot{x} &= Ax + Bu \qquad x \in R^n, u \in R \\ y &= Cx + Du \qquad y \in R \end{aligned} \tag{A.16}$$

is said to be *reachable* if we can find an input $u(t)$ defined on the interval $[0, T]$ that can steer the system from a given final point $x(0) = x_0$ to a desired final point $x(T) = x_f$.

2. The *reachability matrix* for a linear system is given by

$$W_r = \begin{bmatrix} B & AB & \cdots & A^{n-1}B \end{bmatrix}. \tag{A.17}$$

A linear system is reachable if and only if the reachability matrix W_r is invertible (assuming a single intput/single output system). Systems that are not reachable have states that are constrained to have a fixed relationship with each other.

3. *Integral feedback* can be used to provide zero steady state error instead of careful calibration of the gain K_r. An integral feedback controller has the form

$$u = -k_p(x - x_e) - k_i z + k_r r. \tag{A.18}$$

where

$$\dot{z} = y - r \tag{A.19}$$

is the integral error. The gains k_p, k_i and k_r can be found by designing a stabilizing state feedback for the system dynamics augmented by the integrator dynamics.

4. A linear system with dynamics

$$\begin{aligned} \dot{x} &= Ax + Bu \qquad x \in R^n, u \in R \\ y &= Cx + Du \qquad y \in R \end{aligned} \tag{A.20}$$

is said to be *observable* if we can determine the state of the system through measurements of the input $u(t)$ and the output $y(t)$ over a time interval $[0, T]$.

5. The *observability matrix* for a linear system is given by

$$W_o = \begin{pmatrix} C \\ CA \\ \vdots \\ CA^{n-1} \end{pmatrix}. \tag{A.21}$$

A linear system is observable if and only if the observability matrix W_o is full rank. Systems that are not reachable have ”hidden” states that cannot be determined by looking at the inputs and outputs.

6. An *observer* is a dynamical system that estimates the state of another system through measurement of inputs and outputs. For a linear system, the observer given by

$$\frac{d\hat{x}}{dt} = A\hat{x} + Bu + L(y - C\hat{x}) \tag{A.22}$$

generates an estimate of the state that converges to the actual state if $A - LC$ is has eigenvalues with negative real part. If a system is observable, then there exists a an *observer gain* L such that the observer error is governed by a linear differential equation with an arbitrary characteristic polynomial. Hence the eigenvalues of the error dynamics for an observable linear system can be placed arbitrarily through the use of an appropriate observer gain.

7. A discrete time, linear process with noise is given by

$$\begin{aligned} x(k+1) &= Ax(k) + Bu(k) + v(k) \qquad & x \in R^n, u \in R \\ y(k) &= Cx(k) + Du(k) + w(k) \qquad & y \in R \end{aligned} \tag{A.23}$$

where v is a vector, white, Gaussian random process with mean 0, autocovariance R_w, w is a white, Guassian random process with mean 0, variance R_v. We take the initial condition to be random with mean 0 and covariance P_0. The optimal estimator is given by

$$\hat{x}(k+1) = A\hat{x}(k) + Bu(k) + L(y(k) - C\hat{x}(k)) \tag{A.24}$$

where the observer gain satisfies

$$\begin{aligned} P(k+1) &= A^T P(k) A^T + R_v - AP(k)C^T (R_w + CPC^T)^{-1} CP^T(k)A^T \\ P(0) &= P_0 \\ L &= A^T P(k) C^T (R_w + CPC^T)^{-1} \end{aligned} \tag{A.25}$$

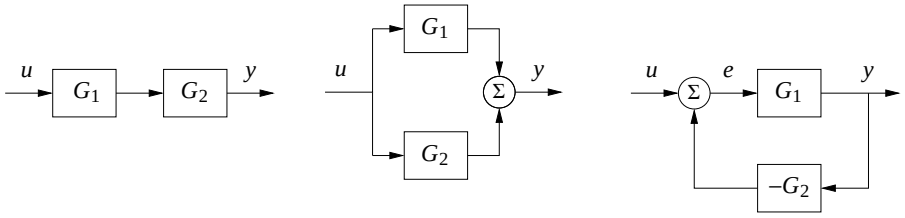
This estimator is an example of a *Kalman filter*.

A.5 Transfer Functions

This chapter introduces the concept of the transfer function, which is a compact description of the input-output relation for a linear system. Combining transfer functions with block diagrams gives a powerful method of dealing with complex systems. The relationship between transfer functions and other system descriptions of dynamics is also discussed.

1. The *frequency response* of a linear system

$$\begin{aligned} \dot{x} &= Ax + Bu \\ y &= Cx + Du \end{aligned} \tag{A.26}$$



(a) $G_{yu} = G_2 G_1$ (b) $G_{yu} = G_1 + G_2$ (c) $G_{yu} = \dfrac{G_1}{1 + G_1 G_2}$

Figure A.2: Interconnections of linear systems. Series (a), parallel (b) and feedback (c) connections are shown. The transfer functions for the composite systems can be derived by algebraic manipulations assuming exponential functions for all signals.

is the response of the system to a sinusoidal input at a given frequency. Due to linearity, the response of a system to a more complicated input can be constructed by decomposing the input into the sum of sines and cosines

$$u(t) = \sum_{k=1}^{\infty} a_k \sin(k\omega t) + b_k \cos(k\omega t). \tag{A.27}$$

2. The *transfer function* for a linear system is given by

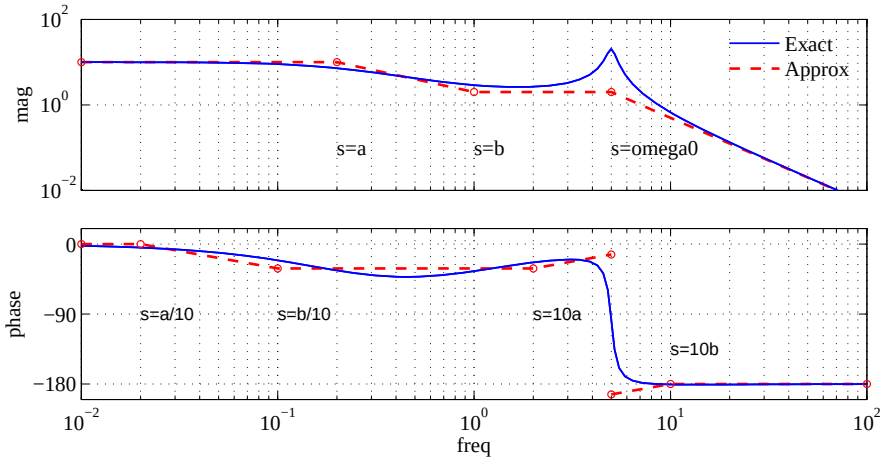
$$G_{yu}(s) = C(sI - A)^{-1}B + D. \tag{A.28}$$

The transfer function represents the steady state response of the system to an exponential input. The transfer function is independent of the choice of coordinates for the state space.

3. The *zero frequency gain* of a system is given by the magnitude of the transfer function at $s = 0$. It represents the ratio of the steady state value of the output with respect to a step input. For a transfer function of the form $G(s) = b(s)/a(s)$, the roots of the polynomial $a(s)$ are called the *poles* of the system and the roots of the polynomial $b(s)$ are called the *zeros* of the system. A pole p is also called a *mode* of the system. The poles correspond to the eigenvalues of the dynamics matrix A and determine the stability of the system. The zeros of a transfer function correspond to exponential signals whose transmission is blocked by the system.

4. Block diagrams that consist of transfer functions can be manipulated using *block diagram algebra*. Figure A.2 gives the transfer functions for some common interconnections of linear systems.

5. A *Bode plot* is a plot of the magnitude and phase of the frequency response:



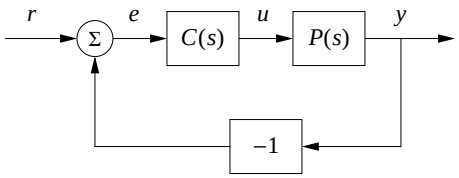
The top plot is the gain curve; the frequency and magnitude are both plotted using a logarithmic scale. The bottom plot is the phase curve and uses a log-linear scale. The dashed lines show straight line approximations of the gain curve and the corresponding phase curve.

6. The transfer function for a system can be determined from experiments by measuring the frequency response and fitting a transfer function to the data. Formally, the transfer function corresponds to the ratio of the Laplace transforms of the output to the input.

A.6 Frequency Domain Analysis

In this chapter we study how how stability and robustness of closed loop systems can be determined by investigating how signals propagate around the feedback loop. The Nyquist stability theorem is a key result that provides a way to analyze stability and introduce measures of degrees of stability.

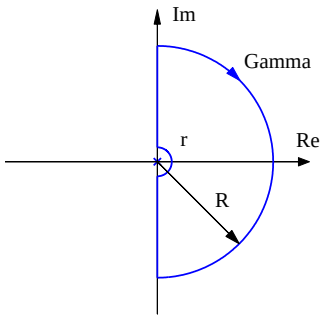
1. The *loop transfer function* of a feedback system represents the transfer function obtained by breaking the feedback loop and computing the resulting transfer function of the open loop system. For a simple feedback system



the loop transfer function is given by $L = PC$

2. The *Nyquist criterion* provides a way to check the stability of a closed loop system by looking at the properties of the loop transfer function. For a stable open loop system, the Nyquist criterion states that the system is stable if the contour of the loop transfer function plotted from $s = -j\infty$ to $s = j\infty$ has no net encirclements of the point $s = -1$ when it is plotted on the complex plane.

3. The general Nyquist criterion uses the image of the loop transfer function applied to the *Nyquist countour*



   The number of unstable poles of the closed loop system is given by the number of open loop unstable poles plus the number of clockwise encirclements of the point $s = -1$.

4. Stability margins describe the robustness of a system to perturbations in the dynamics. We define the *phase crossover frequency*, ω_180 as the smallest frequency where the phase of the loop transfer function is -180° and the *gain crossover frequency*, ω_{gc} as the small frequency where the loop transfer function has unit magnitude. The *gain margin* and *phase margin* are given by

$$g_m = \frac{1}{|L(j\omega_{180})|} \qquad \varphi_m = \pi + \arg L(j\omega_{gc}) \tag{A.29}$$

   These margins describe the the maximum variation in gain and phase in the loop transfer function under which the system remains stable. Two other margins are the *stability margin*, which is the shortest distance frmo the Nyquist curve to the critical point $s = -1$, and the *delay margin*, which is the smallest time delay required to make the system unstable.

5. *Bode's relations* relate the gain and phase of a transfer function with no poles or zeros in the right half plane. They show that

$$\arg G(j\omega_0) \approx \frac{\pi}{2} \frac{d \log |G(j\omega)|}{d \log \omega}. \tag{A.30}$$

A *non-minimum phase* sytem is one for which there is more phase lag than the amount given by Bode's relations. Systems with right have plane poles or zeros are non-minimum phase.

6. The *gain* of an input/output system is defined as

$$\gamma = \sup_{u \in \mathcal{U}} \frac{\|y\|}{\|u\|}, \tag{A.31}$$

where sup is the supremum. The *small gain theorem* states that if two systems with gains γ_1 and γ_2 are connected in a feedback loop, then the closed loop system is stable if $\gamma_1\gamma_2$.

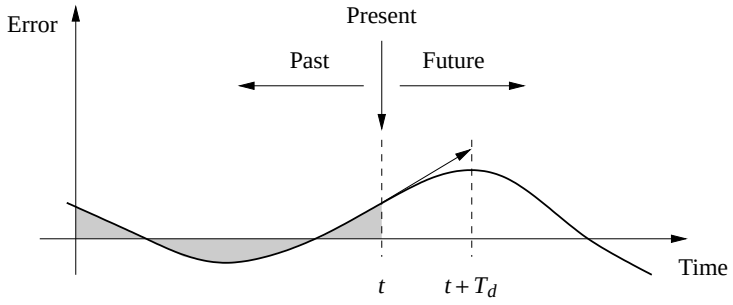
A.7 PID Control

This chapter describes the use of proportional integral derivative (PID) feedback for control systems design. We discuss the basic concepts behind PID control and the methods for choosing the PID gains.

1. The basic PID controller as the form

$$u(t) = k_p e(t) + k_i \int_0^t e(\tau)\, d\tau + k_d \frac{de}{dt}, \tag{A.32}$$

where u is the control signal and e is the control error. The control signal is thus a sum of three terms: a proportional term that is proportional to the error, an integral term that is proportional to the integral of the error, and a derivative term that is proportional to the derivative of the error.



2. *Integral action* guarantees that the process output agrees with the reference in steady state and provides an alternative to including a feedforward term for tracking a constant reference input. Integral action can be implemented using *automatic reset*, where the output of a proportional controller is fed back to its input through a low pass filter:

$$u = k_p e + \frac{1}{1 + sT_i} u, \tag{A.33}$$

3. *Derivative action* provides a method for predictive action. The input-output relation of a controller with proportional and derivative action is

$$u = k_p e + k_d \frac{de}{dt} = k(e + T_d \frac{de}{dt}), \tag{A.34}$$

where $T_d = k_d/k_p$ is the derivative time constant. The action of a controller with proportional and derivative action can be interpreted as if the control is made proportional to the predicted process output, where the prediction is made by extrapolating the error T_d time units into the future using the tangent to the error curve.

A.8 Limits of Performance

In this chapter we continue to explore the use of frequency domain techniques for design of feedback systems. We begin with a more thorough description of the performance specifications for controls systems, and then introduce the concept of ”loop shaping” as a mechanism for designing controllers in the frequency domain. We also introduce some fundamental limitations to performance for systems with right half plane poles and zeros.

1. The primary transfer functions that define the input/output characteristics of the system are called the *Gang of Six*:

$$\begin{aligned} TF &= \frac{PCF}{1+PC}, & T &= \frac{PC}{1+PC}, & PS &= \frac{P}{1+PC}, \\ CFS &= \frac{CF}{1+PC}, & CS &= \frac{C}{1+PC}, & S &= \frac{1}{1+PC}. \end{aligned} \tag{A.35}$$

The transfer functions in the first column give the response of the process output and control signal to the reference signal. The second column gives the response of the control variable to the load disturbance and the noise, and the final column gives the response of the process output to those two inputs. When $F(s) = 1$, the system is said to have pure error feedback and the relevant input/output transfer functions are given by the *Gang of Four*, given by the transfer functions in the right two columns.

2. The performance of a system can be given in terms of the characteristics of the frequency response between an input and output. A *resonant peak* is a maximum of the gain, and the peak frequency is the corresponding frequency.

3. The *sensitivity function* $S = 1/(1 + PC)$ describes how disturbances are attenuated by closing the feedback loop. Disturbances with frequencies such

that $|S(i\omega)| < 1$ are attenuated, but disturbances with frequencies such that $|S(i\omega)| > 1$ are amplified by feedback. The maximum sensitivity M_s, which occurs at the frequency ω_{ms}, is a measure of the largest amplification of the disturbances. The *complementary sensitivity function* $T = PC/(1 + PC)$ describes how well the controller tracks a references signal. The *maximum complementary sensitivity*, M_t, which occurs at the frequency ω_{mt}, is the peak value of the magnitude of the complementary sensitivity function. It provides the maximum amplification from the reference signal to the output signal.

4. Feedback control systems have a number of fundamental limits, usually exacerbated by the presence of right half plane poles and zeros. For systems with right half plane poles or zeros, we can decompose the process dynamics into a minimum phase transfer function (no right half plane poles or zeros) and an all pass transfer function (gain = 1):

$$P(s) = P_{mp}(s)P_{ap}(s), \tag{A.36}$$

5. Another fundamental limit is given by *Bode's integral formula*, which states that for systems with a loop transfer function that goes to zero faster than $1/s$ as $s \to \infty$, the sensitivity function must satisfy

$$\int_0^\infty \log|S(i\omega)|\,d\omega = \int_0^\infty \log\frac{1}{|1 + L(i\omega)|}\,d\omega = \pi\sum p_k, \tag{A.37}$$

where p_k are the poles in the right half-plane. This conservation law shows that to get lower sensitivity in one frequency range, we must get higher sensitivity in some other region. An analogous formula exists for the complementary sensitivity function in the presence of right half plane zeros.

A.9 Robust Performance

This chapter focuses on the analysis of robustness of feedback systems. We consider the stability and performance of systems who process dynamics are uncertain and derive fundamental limits for robust stability and performance. We also discuss how to design controllers to achieve robust performance.

1. Uncertainty can enter a model in many forms. *Parametric uncertainty* occurs when the values of the parameters in the model are not precisely known or may vary. *Unmodeled dynamics* are a more general class of uncertainty in which some portions of the systems behavior are not included in the model, either due to lack of knowledge or simplicity. Unmodeled dynamics can be taken into consideration by incorporating an uncertainty block with bounded input/output response. Common types of unmodeled dynamics include *additive uncertainty*, *multiplicative uncertainty* and *feedback uncertainty*.

2. The *Vinnicombe metric* (or ν-gap metric) provides a measure of the distance between two transfer functions. It is defined as

$$\delta_\nu(P_1, P_2) = \begin{cases} d(P_1, P_2), & \text{if } (P_1, P_2) \in C \\ 1, & \text{otherwise,} \end{cases} \tag{A.38}$$

where $d(P_1, P_2)$ is a distance measure between the two transfer function

$$d(P_1, P_2) = \sup_\omega \frac{|P_1(i\omega) - P_2(i\omega)|}{\sqrt{(1 + |P_1(i\omega)|^2)(1 + |P_2(i\omega)|^2)}}, \tag{A.39}$$

and C is the set of all pairs (P_1, P_2) such that the functions $f_1 = 1 + P_1(s)P_1(-s)$ and $f_2 = 1 + P_2(s)P_1(-s)$ have the same number of zeros in the right half-plane

3. Robust stability can be determined through the use of the Nyquist plot. The *stability margin* s_m, defined as the shortest distanced from -1 to the Nyquist curve, provides a measure of robustness. For an additive perturbation $\Delta(s)$, the system is robustly stable if

$$|\Delta| < \left|\frac{1 + PC}{C}\right| \qquad \text{or} \qquad |\delta| = \left|\frac{\Delta}{P}\right| < \frac{1}{|T|}. \tag{A.40}$$

This condition can be derived using the *small gain theorem* and allows us to reason about uncertainty without exact knowledge of the process perturbations.

4. In addition to stability, uncertainty can also affect the performance of a system. For additive uncertainty, the load response satisfies

$$\frac{dG_{yd}}{G_{yd}} = S\frac{dP}{P}. \tag{A.41}$$

The response to load disturbances is thus insensitive to process variations for frequencies where the magnitude of the sensitivity function $|S(i\omega)|$ is small. Similarly, the response of the controller to noise in the presence of additive uncertainty satisfies

$$\frac{dG_{un}}{G_{un}} = T\frac{dP}{P}, \tag{A.42}$$

indicating that the controller is insensitive to noise when the complementary sensitivity is small. Control design in the presence of uncertainty can be done by using the Gang of Four to insure that the appropriate sensitivity functions are all well behaved.

Bibliography

[1] K. J. Åström and R. M. Murray. *Feedback Systems: An Introduction for Scientists and Engineers*. Princeton University Press, 2008. Available at http://www.cds.caltech.edu/~murray/amwiki.

[2] B. Alberts, D. Bray, J. Lewis, M. Raff, K. Roberts, and J. D. Watson. *The Molecular Biology of the Cell*. Garland Science, fifth edition edition, 2008.

[3] U. Alon. *An introduction to systems biology. Design principles of biological circuits*. Chapman-Hall, 2007.

[4] W. Arber and S. Linn. DNA modification and restriction. *Annual Review of Biochemistry*, 38:467–500, 1969.

[5] M. R. Atkinson, M. A. Savageau, J. T. Meyers, and A. J. Ninfa. Development of genetic circuitry exhibiting toggle switch or oscillatory behavior in *Escherichia coli*. *Cell*, pages 597–607, 2003.

[6] D. W. Austin, M. S. Allen, J. M. McCollum, R. D. Dar, J. R. Wilgus, G. S. Sayler, N. F. Samatova, C. D. Cox, and M. L. Simpson. Gene network shaping of inherent noise spectra. *Nature*, 2076:608–611, 2006.

[7] D. Baker, G. Church, J. Collins, D. Endy, J. Jacobson, J. Keasling, P. Modrich, C. Smolke, and R. Weiss. ENGINEERING LIFE: Building a FAB for biology. *Scientific American*, June 2006.

[8] N Barkai and S Leibler. Robustness in simple biochemical networks. *Nature*, 387(6636):913–7, 1997.

[9] A. Becskei and L. Serrano. Engineering stability in gene networks by autoregulation. *Nature*, 405:590–593, 2000.

[10] D. Bell-Pedersen, V. M. Cassone, D. J. Earnest, S. S. Golden, P. E. Hardin, T. L. Thomas, and M. J. Zoran. Circadian rhythms from multiple oscillators: lessons from diverse organisms. *Nature Reviews Genetics*, 6(7):544, 2005.

[11] BioNumbers: The database of useful biological numbers. http://bionumbers.org, 2012.

[12] L Bleris, Z. Xie, D. Glass, A. Adadey, E. Sontag, and Y. Benenson. Synthetic incoherent feedforward circuits show adaptation to the amount of their genetic template. *Molecular Systems Biology*, 7:519, 2011.

[13] H. Bremer and P. Dennis. Modulation of chemical composition and other parameters of the cell by growth rate. *In: Escherichia coli and Salmonella: Cellular and Molecular Biology (edited by Neidhart F. C. et al.), ASM Press, Washington DC*, 183:1553–1569, 1996.

[14] H Bremer, P Dennis, and M Ehrenberg. Free rna polymerase and modeling global transcription in *escherichia coli*. *Biochimie*, 85:597–609, 2003.

[15] B. Canton, A. Labno, and D. Endy. Refinement and standardization of synthetic biological parts and devices. *Nature Biotechnology*, 26(7):787–93, 2008.

[16] M. Chalfie, Y. Tu, G. Euskirchen, W. Ward, and D. Prasher. Green fluorescent protein as a marker for gene expression. *Science*, 263(5148):802–805, 1994.

[17] A. J. Courey. *Mechanisms in Transcriptional Regulation*. Wiley-Blackwell, 2008.

[18] R. S. III Cox, M. G. Surette, and M. B. Elowitz. Programming gene expression with combinatorial promoters. *Mol Syst Biol*, page 3:145, 2007.

[19] D. Del Vecchio, A. J. Ninfa, and E. D. Sontag. Modular cell biology: Retroactivity and insulation. *Nature/EMBO Molecular Systems Biology*, 4:161, 2008.

[20] L. N. M. Duysens and J. Amesz. Fluorescence spectrophotometry of reduced phosphopyridine nucleotide in intact cells in the near-ultraviolet and visible region. *Biochim. Biophys. Acta*, 24:19–26, 1957.

[21] H. El-Samad, J. P. Goff, and M. Khammash. Calcium homeostasis and parturient hypocalcemia: An integral feedback perspective. *J. Theoret. Biol.*, 214:17–29, 2002.

[22] S. P. Ellner and J. Guckenheimer. *Dynamic Models in Biology*. Princeton University Press, Princeton, NJ, 2005.

[23] M. B. Elowitz and S. Leibler. A synthetic oscillatory network of transcriptional regulators. *Nature*, 403(6767):335–338, 2000.

[24] Michael B Elowitz, Arnold J Levine, Eric D Siggia, and Peter S Swain. Stochastic gene expression in a single cell. *Science (New York, NY)*, 297(5584):1183–1186, 2002.

[25] D. Endy. Foundations for engineering biology. *Nature*, 438:449–452, 2005.

[26] B. Friedland. *Control System Design: An Introduction to State Space Methods*. Dover, New York, 2004.

[27] T.S. Gardner, C.R. Cantor, and J.J. Collins. Construction of the genetic toggle switch in *Escherichia Coli*. *Nature*, page 339342, 2000.

[28] Daniel G. Gibson, John I. Glass, Carole Lartigue, Vladimir N. Noskov, Ray-Yuan Chuang, Mikkel A. Algire, Gwynedd A. Benders, Michael G. Montague, Li Ma, Monzia M. Moodie, Chuck Merryman, Sanjay Vashee, Radha Krishnakumar, Nacyra Assad-Garcia, Cynthia Andrews-Pfannkoch, Evgeniya A. Denisova, Lei Young, Zhi-Qing Qi, Thomas H. Segall-Shapiro, Christopher H. Calvey, Prashanth P. Parmar, Clyde A. Hutchison, Hamilton O. Smith, and J. Craig Venter. Creation of a Bacterial Cell Controlled by a Chemically Synthesized Genome. *Science*, 329(5987):52–56, 2010.

[29] D. T. Gillespie. *Markov Processes: An Introduction For Physical Scientists*. Academic Press, 1976.

[30] D. T. Gillespie. Exact stochastic simulation of coupled chemical reactions. *Journal of Physical Chemistry*, 81(25):2340–2361, 1977.

[31] D. T. Gillespie. A rigorous derivation of the chemical master equation. *Physica A*, 188:404–425, 1992.

[32] L. Goentoro, O. Shoval, M. W. Kirschner, and U. Alon. The incoherent feedforward loop can provide fold-change detection in gene regulation. *Molecular Cell*, 36:894–899, 2009.

[33] A. Goldbeter and D. E. Koshland. An amplified sensitivity arising from covalent modification in biological systems. *PNAS*, pages 6840–6844, 1981.

[34] J. Greenblatt, J. R. Nodwell, and S. W. Mason. Transcriptional antitermination. *Nature*, 364(6436):401–406, 1993.

[35] J. Greenblatt, J. R. Nodwell, and S. W. Mason. Transcriptional antitermination. *Nature*, 364(6436):401–406, 1993.

[36] I.L. Grigiriva, N.J. Phleger, V.K. Mutalik, and C.A. Gross. Insights into transcriptional regulation and σ competition from an equilibrium model of RNA polymerase binding to DNA. *PNAS*, 103(14):5332–5337, 2006.

[37] J. Guckenheimer and P. Holmes. *Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields*. Springer, 1983.

[38] R. Heinrich, B. G. Neel, and T. A. Rapoport. Mathematical models of protein kinase signal transduction. *Molecular Cell*, 9:957–970, 2002.

[39] B. Hess, A. Boiteux, and J. Kruger. Cooperation of glycolytic enzymes. *Adv. Enzyme Regul*, 7:149–167, 1969.

[40] Andreas Hilfinger and Johan Paulsson. Separating intrinsic from extrinsic fluctuations in dynamic biological systems. *Proceedings of the National Academy of Sciences*, 108(29):12167–12172, 2011.

[41] C. F. Huang and J. E. Ferrell. Ultrasensitivity in the mitogen-activated proteinkinase cascade. *Proc. Natl. Acad. Sci.*, 93(19):10078–10083, 1996.

[42] T. P. Hughes. *Elmer Sperry: Inventor and Engineer*. John Hopkins University Press, Baltimore, MD, 1993.

[43] B. Ingalls. A frequency domain approach to sensitivity analysis of biochemical networks. *Journal of Physical Chemistry B-Condensed Phase*, 108(3):143–152, 2004.

[44] A Ishihama. Functional modulation of *e. coli* rna polymerase. *Ann. Rev. Microbiol*, 54:499–518, 2000.

[45] F. Jacob and J. Monod. Genetic regulatory mechanisms in the synthesis of proteins. *J. Mol. Biol.*, 3:318–56, 1961.

[46] S. Jayanthi, K. Nilgiriwala, and D. Del Vecchio. Retroactivity controls the temporal dynamics of gene transcription. *ACS Synthetic Biology*, DOI: 10.1021/sb300098w, 2013.

[47] K. F. Jensen and S. Pedersenz. Metabolic growth rate control in*escherichia coli* may be a consequence of subsaturation of the macromolecular biosynthetic apparatus with substrates and catalytic components. *MICROBIOLOGICAL REVIEWS*, 54(2):89–100, 1990.

[48] P. Jiang, A. C. Ventura, S. D. Merajver, E. D. Sontag, A. J. Ninfa, and D. Del Vecchio. Load-induced modulation of signal transduction networks. *Science Signaling*, 4(194):ra67, 2011.

[49] N. G. Van Kampen. *Stochastic Processes in Physics and Chemistry*. Elsevier, 1992.

[50] A. S. Khalil and J. J. Collins. Synthetic biology: applications come of age. *Nature Reviews Genetics*, 11(5):367, 2010.

[51] H. K. Khalil. *Nonlinear Systems*. Macmillan, 1992.

[52] E. Klipp, W. Liebermeister, C. Wierling, A. Kowald, H. Lehrach, and R. Herwig. *Systems Biology: A Textbook*. Wiley-VCH, 2009.

[53] P. Kundur. *Power System Stability and Control*. McGraw-Hill, New York, 1993.

[54] M. T. Laub, L. Shapiro, and H. H. McAdams. Systems biology of *caulobacter*. *Annual Review of Genetics*, 51:429–441, 2007.

[55] J.-C. Leloup and A. Goldbeter. A molecular explanation for the long-term supression of circadian rhythms by a single light pulse. *American Journal of Physiology*, 280:1206–1212, 2001.

[56] J. J. Lemke, P. Sanchez-Vazquez, H. L. Burgos, G. Hedberg, W. Ross, and R. L. Gourse. Direct regulation of *Escherichia coli* ribosomal protein promoters by the transcription factors ppGpp and DksA. *PNAS*, pages 1–6, 2012.

[57] W. Lohmiller and J. J. E. Slotine. On contraction analysis for non-linear systems. *Automatica*, 34:683–696, 1998.

[58] H. Madhani. *From a to alpha: Yeast as a Model for Cellular Differentiation*. CSHL Press, 2007.

[59] J. Mallet-Paret and H.L. Smith. The Poincaré-Bendixson theorem for monotone cyclic feedback systems. *J. of Dynamics and Differential Equations.*, 2:367–421, 1990.

[60] J. E. Marsden and M. J. Hoffman. *Elementary Classical Analysis*. Freeman, 2000.

[61] S. Marsigliante, M. G. Elia, B. Di Jeso, S. Greco, A. Muscella, and C. Storelli. Increase of [ca(2+)](i) via activation of atp receptors in pc-cl3 rat thyroid cell line. *Cell. Signal*, 14:61–67, 2002.

[62] H. H. McAdams and Arkin A. Stochastic mechanisms in gene expression. *PNAS*, 94:814–819, 1997.

[63] C. R. McClung. Plant circadian rhythms. *Plant Cell*, 18:792–803, 2006.

[64] M. W. McFarland, editor. *The Papers of Wilbur and Orville Wright*. McGraw-Hill, New York, 1953.

[65] P. Miller and X. J. Wang. Inhibitory control by an integral feedback signal in prefrontal cortex: A model of discrimination between sequential stimuli. *PNAS*, 103:201–206, 2006.

[66] C. J. Morton-Firth, T. S. Shimizu, and D. Bray. A free-energy-based stochastic simulation of the tar receptor complex. *Journal of Molecular Biology*, 286(4):1059–74, 1999.

[67] J. D. Murray. *Mathematical Biology*, Vols. I and II. Springer-Verlag, New York, 3rd edition, 2004.

[68] R. M. Murray. *Optimization-Based Control*. http://www.cds.caltech.edu/~murray/amwiki/OBC, Retrieved 20 December 2009.

[69] C. J. Myers. *Engineering Genetic Circuits*. Chapman and Hall/CRC Press, 2009.

[70] T. Nagashima, H. Shimodaira, K. Ide, T. Nakakuki, Y. Tani, K. Takahashi, N. Yumoto, and M. Hatakeyama. Quantitative transcriptional control of erbb receptor signaling undergoes graded to biphasic response for cell differentiation. *J. Biol. Chem.*, 282:40454056, 2007.

[71] R. Nesher and E. Cerasi. Modeling phasic insulin release: Immediate and time-dependent effects of glucose. *Diabetes*, 51:53–59, 2002.

[72] R. Phillips, J. Kondev, and J. Theriot. *Physical Biology of the Cell*. Garland Science, 2008.

[73] M. Ptashne. *A genetic switch*. Blackwell Science, Inc., 1992.

[74] P. E. M. Purnick and R. Weiss. The second wave of synthetic biology: from modules to systems. *Nature Reviews Molecular Cell Biology*, 10(6):410–422, 2009.

[75] E. K. Pye. Periodicities in intermediary metabolism. *Biochronometry, National Acad. Sci.*, 1971.

[76] L. Qiao, R. B. Nachbar, I. G. Kevrekidis, and S. Y. Shvartsman. Bistability and oscillations in the Huang-Ferrell model of MAPK signaling. *PLoS Computational Biology*, 3:e184, 2007.

[77] C. V. Rao, J. R. Kirby, and A. P. Arkin. Design and diversity in bacterial chemotaxis: A comparative study in escherichia coli and bacillus subtilis. *PLoS Biology*, 2(2):239–252, 2004.

[78] J. M. Rohwer, N. D. Meadow, S. Roseman, H. V. Westerhoff, and P. W. Postma. Understanding glucose transport by the bacterial phosphoenolpyruvate: glucose phosphotransferase system on the basis of kinetic measurements in vitro. *The Journal of biological chemistry*, 275(45):34909–34921, November 2000.

[79] H. M. Sauro and B. N. Kholodenko. Quantitative analysis of signaling networks. *Progress in Biophysics & Molecular Biology*, 86:5–43, 2004.

[80] D. E. Seborg, T. F. Edgar, and D. A. Mellichamp. *Process Dynamics and Control*. Wiley, Hoboken, NJ, 2nd edition, 2004.

[81] Thomas S Shimizu, Yuhai Tu, and Howard C Berg. A modular gradient-sensing network for chemotaxis in Escherichia coli revealed by responses to time-varying stimuli. *Molecular Systems Biology*, 6:382, 2010.

[82] O. Shimomura, F. Johnson, and Y. Saiga. Extraction, purification and properties of aequorin, a bioluminescent protein from the luminous hydromedusan, Aequorea. *J Cell Comp Physiol*, 59(3):223–239, 1962.

[83] O. Shoval, U. Alon, and E. Sontag. Symmetry invariance for adapting biological systems. *SIAM J. APPLIED DYNAMICAL SYSTEMS*, 10:857886, 2011.

[84] E. D. Sontag. *Mathematical Control Theory: Deterministic Finite Dimensional Systems*. Springer, New York, 2nd edition, 1998.

[85] E.D. Sontag. Remarks on feedforward circuits, adaptation, and pulse memory. *IET Systems Biology*, 4:39–51, 2010.

[86] J. Stricker, S. Cookson, M. R. Bennett, W. H. Mather, L. S. Tsimring, and J. Hasty. A fast, robust and tunable synthetic gene oscillator. *Nature*, 456(7221):516–519, 2008.

[87] J. Stülke and W. Hillen. Coupling physiology and gene regulation in bacteria: the phosphotransferase sugar uptake system delivers the signals. *Die Naturwissenschaften*, 85(12):583–592, December 1998.

[88] Peter S Swain, Michael B Elowitz, and Eric D Siggia. Intrinsic and extrinsic contributions to stochasticity in gene expression. *Proceedings of the National Academy of Sciences of the United States of America*, 99(20):12795–12800, 2002.

[89] J. Tsang, J. Zhu, and A. van Oudenaarden. Microrna-mediated feedback and feedforward loops are recurrent network motifs in mammals. *Mol. Cell*, 26:753–767, 2007.

[90] K. V. Venkatesh, S. Bhartiya, and A. Ruhela. Mulitple feedback loops are key to a robust dynamic performance of tryptophan regulation in *Escherichia coli*. *FEBS Letters*, 563:234–240, 2004.

[91] A. C. Ventura, P. Jiang, L. Van Wassenhove, D. Del Vecchio, S. D. Merajver, and A. J. Ninfa. The signaling properties of a covalent modification cycle are altered by a downstream target. *Proc. Natl. Acad. Sci. USA*, 107(22):10032–10037, 2010.

[92] O. Venturelli, H. El-Samad, and R. M. Murray. Dual positive feedback loops compensate for generating robust bistability in the gal regulatory network. In preparation, 2012.

[93] O. S. Venturelli, H. El-Samad, and R. M. Murray. Synergistic dual positive feedback loops established by molecular sequestration generate robust bimodal response. *Proc. of the National Academy of Sciences*, 109(48):E3324–33, 2012.

[94] L. Villa-Komaroff, A. Efstratiadis, S. Broome, P. Lomedico, R. Tizard, S. P. Naber, W. L. Chick, and W. Gilbert. A bacterial clone synthesizing proinsulin. *Proc. Natl. Acad. Sci. U.S.A.*, 75(8):372731, 1978.

[95] C. A. Voigt. Genetic parts to program bacteria. *Current Opinions in Biotechnology*, 17(5):548–557, 2006.

[96] S. Wiggins. *Introduction to Applied Nonlinear Dynamical Systems and Chaos*. Springer, 2003.

[97] L. Yang and P. A. Iglesias. Positive feedback may cause the biphasic response observed in the chemoattractant-induced response of dictyostelium cells. *Systems Control Lett.*, 55:329 337, 2006.

[98] T.-M. Yi, Y. Huang, M. I. Simon, and J. C. Doyle. Robust perfect adaptation in bacterial chemotaxis through integral feedback control. *Proc. of the National Academy of Sciences*, 97(9):4649–4653, 2000.

[99] N. Yildirim and M. C. Mackey. Feedback regulation in the lactose operon: A mathematical modeling study and comparison with experimental data. *Biophysical Journal*, 84(5):2841–2851, 2003.

[100] N. Yildirim, M. Santillan, D. Horike, and M. C. Mackey. Dynamics and bistability in a reduced model of the lac operon. *Chaos*, 14(2):279–292, 2004.

Index