Molecular Systems Biology and Control*

Eduardo D. Sontag
Department of Mathematics and BioMaPS Institute for Quantitative Biology
Rutgers University, New Brunswick, NJ 08903, USA
E-mail: sontag@math.rutgers.edu

Abstract

This paper, prepared for a tutorial at the 2005 IEEE Conference on Decision and Control, presents an introduction to molecular systems biology and some associated problems in control theory. It provides an introduction to basic biological concepts, describes several questions in dynamics and control that arise in the field, and argues that new theoretical problems arise naturally in this context. A final section focuses on the combined use of graph-theoretic, qualitative knowledge about monotone building-blocks and steady-state step responses for components.

1 Introduction

Within the last few years, the field of "molecular systems biology" has taken shape, having as its goal the unraveling of the basic dynamic processes, feedback control loops, and signal processing mechanisms underlying life. Leading biologists have recognized that new systems-level knowledge is urgently required in order to conceptualize and organize the revolutionary developments taking place in the biological sciences, and new academic departments and educational programmes are being established at major universities, particularly in Europe and in the United States.

The studies of dynamics, feedback, and signal processing in engineering and in biology have long been intertwined, for example in the fields of biological and biomedical engineering. But our community has also actively participated in the study of biological control systems in their own right, independent of such application areas. Indeed, one of the founders of our field, Norbert Wiener, developed many of the ideas of feedback and filtering in the early 1940s in collaboration with the Harvard physiologist Arturo Rosenblueth, who was, in turn, heavily influenced by the work of his colleague Walter Cannon, who coined the term *homeostasis* in 1932 to refer to feedback mechanisms for set-point regulation in living organisms. Wiener viewed his study of *cybernetics* as a unifying theme in engineering and biology. Rudolf Kalman often used biological analogies in his discussion of control systems theory, and so did many other early researchers. Balthazar van der Pol, the Dutch electrical engineer whose oscillator models of vacuum tubes are a routine example in the theory of limit cycles, was motivated by models of the human heart and an interest in arrhythmias. In parallel, and for at least as long, mathematical biologists have been developing quantitative theories of physiological regulation, metabolic pathways, insulin control, heart electrical patterns, neural and circadian oscillations, and so forth.

So, one may ask, why the sudden resurgence of interest? The answer surely involves a combination of many factors. Bioinformatics has been tremendously successful in facilitating the sequencing of human, animal, plant, bacterial, and other genomes, as well as in protein structure prediction. Nontrivial ideas and algorithms from discrete mathematics, probability and statistics, theoretical computer science, and even partially observed stochastic systems (Hidden Markov Models), embedded in user-friendly software, are now indispensable tools

*To appear in *European J. of Control* **11**(2005): 1-40.

of the working biologist and pharmaceutical researcher. Thus, many biologists have come to accept and value the use of mathematical tools. On the other hand, new data collection and measurement approaches, themselves based upon sophisticated engineering, make possible the simultaneous monitoring of the activity of thousands of genes and the concentrations of proteins and metabolites, thus allowing for the study of microscopic dynamic interactions among cellular components, and making a systems-level view of cells particularly natural. The huge amounts of data being generated by genomics and proteomics require new theoretical approaches to interpretation and organization. Medical advances also drive this new emphasis. Many in the pharmaceutical industry have come to the realization that only by understanding cells as a whole can one identify novel targets for new drugs, and understand their systemic effects; gene therapies will depend on a more global understanding of dynamic interactions among genes and their cellular environment. Finally, and at a somewhat more philosophical level, there also is the fact that current experimental methods permit making *falsifiable predictions*, bringing modern biology closer to physics and chemistry as a science.

While classical theoretical biology dealt largely with ecology, or with whole organisms, biologists can now test hypotheses in a precisely targeted fashion. For example, if a mathematical model predicts that a certain mutation will make fruit flies grow a leg instead of antennae on their heads, the mutation can be carried out and the results observed.

Control and systems theory have much to offer to biology. But, conversely, one may look forward to technologies inspired by biological research: evolution has resulted in systems that are highly fault-tolerant, nonlinear, feedback-rich, and truly hybrid —in the sense that the digital information encoded in DNA controls chemical concentrations in cells. Advances in genomic research are continually adding to detailed knowledge of such systems' architecture and operation, and one may reasonably argue that they will constitute a rich source of inspiration for innovative solutions to problems of control and communication engineering, as well as sensor and actuator design and integration.

This paper is organized as follows. The next section —which may be skipped by those readers not interested in, or already knowledgeable about, the chemistry and biology background— provides an introduction to many of the basic concepts of molecular biology. (For a serious study of the subject, a recommended starting point is the book [2].) Next, the paper describes some of the central questions in dynamics and control that arise in the field. It is argued that some of these questions differ in an essential manner from similar-sounding questions in engineering applications, thus leading one to entirely new theoretical control and systems theory problems. The final section focuses on a topic in which the author has recently worked, namely, the combination of network-like, qualitative knowledge, with a comparatively small amount of quantitative data, in order to help characterize global behavior. This approach is based upon decompositions into well-behaved building-blocks —monotone subsystems— and the use of input/output data —steady-state step responses— for these components.

2 Molecular Cell Biology

The fundamental unit of life is the cell (Figure 1). Organisms may consist of just one cell or they may be multicellular; the latter type are typically organized into tissues, which are groups of similar cells arranged so as to perform a specific function. (For example, humans have on the order of 10^{14} cells organized into roughly 200 tissues.)

One may view cell life as a collection of "wireless networks" of interactions among proteins, RNA, DNA, and small molecules involved in signaling and energy transfer. These networks process environmental signals, induce appropriate cellular responses, and sequence internal events such as gene expression, thus allowing cells and entire organisms to perform their basic functions. These control and communication networks can be relatively simple, such as the *two-component systems* found mainly in bacteria, which are cascades connecting sensors (proteins in the cell membrane, which detect outside signals) to actuators (typically transcription factors, which direct the expression of a gene), cf. [94]. Or they may be incredibly sophisticated, as in higher

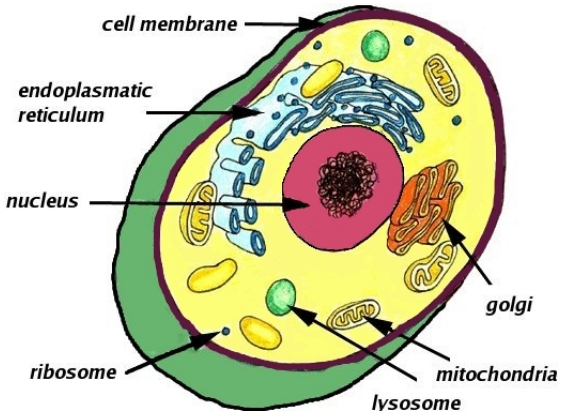

Figure 1: An eukaryotic cell

organisms, involving multiple *signal transduction pathways* in which information is relayed among enzymes through chemical reactions (for instance, phosphorylation).

In addition to their own needs for survival and reproduction, cells in multicellular organisms need additional levels of complexity in order to enable communication among cells and overall regulation, as well as to direct differentiation from a single fertilized egg into the various tissues in an individual member of a species. We will focus on intracellular pathways, but these other aspects are no less exciting areas of study.

Before providing more details and examples, let us step back and review some of the basic concepts and terminology.

2.1 Prokaryotes, Eukaryotes, Archaea, and Viruses

At the highest level, biologists classify life forms into prokaryotes, eukaryotes, and archaea. *Prokaryotes* are organisms whose cells do not have a nucleus nor other well-defined compartments; their genetic information is stored in chromosomes –typically circular– as well as in smaller circular DNA molecules called plasmids. *Eukaryotes* have cells with organized compartments; their genetic material is stored in chromosomes –typically linear– that lie in the nucleus. Most prokaryotes, with few exceptions, are unicellular, and most are bacteria. Eukaryotes might be unicellular (e.g., yeast) or multicellular (e.g., plants and animals). *Archaea* were proposed as a third life form in the mid-1970s, and they share many characteristics with both prokaryotes and eukaryotes.

Eukaryotic cells are enclosed in a *plasma membrane*, which is made up of lipids and also contains proteins and carbohydrates, and acts as a protective barrier and gatekeeper, permitting only selected chemicals to enter and leave the cell. (In addition to membranes, plant cells also have a rigid cell wall.) Their interior is called the cytoplasm, and many types of organelles —specialized compartments— populate the cell (mitochondria, responsible for energy production through metabolism, and containing a very small amount of DNA; chloroplasts for photosynthesis; ribosomes, responsible for protein synthesis, and made up themselves of proteins and RNAs; endoplasmic reticulum; and so forth). The cytoskeleton, made up of microtubules and filaments, gives shape to the cell and plays a role in intracell substance transport. Prokaryotic cells, on the other hand, are surrounded by a membrane and cell wall, but do not contain the usual organelles.

Viruses consist of protein-coated DNA or RNA, and are not usually classified as living organisms, because they cannot reproduce by themselves, but rather require the machinery of a host cell in order to replicate. In particular, bacteriophages are viruses that infect bacteria.

2.2 Genomics and Proteomics

Research in molecular biology, genomics, and proteomics has produced, and will continue to produce, a wealth of data describing the elementary components of intracellular networks as well as detailed mappings of their pathways and environmental conditions required for activation.

2.2.1 DNA and Genes

The *genome*, that is to say, the genetic information of an individual, is encoded in double-stranded *deoxyribonucleic acid (DNA)* molecules, which are arranged into chromosomes. It may be viewed as a "parts list" which describes all the proteins that are potentially present in every cell of a given organism. Genomics research has as its objective the complete decoding of this information, both the parts common for a species as a whole and the cataloging of differences among individual members.

The key paradigm of molecular biology: "DNA makes RNA, RNA makes protein, and proteins make the cell" is called the *central dogma of molecular biology* (Crick, 1958).[1] A separate process, *replication*, occurs more rarely, and only when a cell is ready to divide (S phase of mitosis, in eukaryotes), and results in the duplication of the DNA, one copy to be part of each of the two daughter cells. See Figure 2. The term *gene*

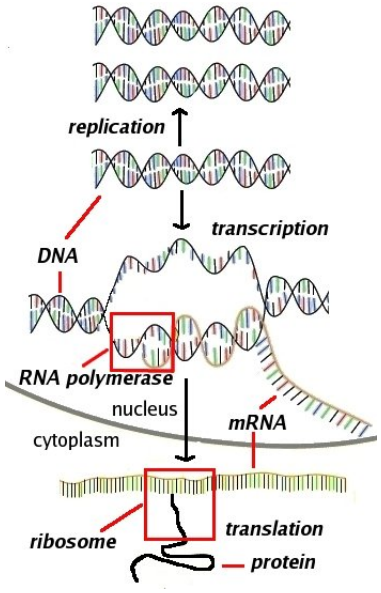

Figure 2: Central dogma of molecular biology

expression refers to the process by which genetic information gets ultimately transformed into working proteins. The main steps are transcription from DNA to RNA, translation from RNA to linear amino acid sequences, and folding of these into functional proteins, but several intermediate editing steps usually take place as well. (Sometimes the term "gene expression" is used only for the transcription part of this process.) At any given time, and in any given cell of an organism, thousands of genes and their products (RNA, proteins) actively participate in an orchestrated manner.

[1] Recent work is forcing a rethinking of the roles of RNA and proteins. For example, prions appear to take advantage of a direct mechanism for protein replication: when a prion infects an organism, it interacts with wild-type –that is to say, normal– proteins, causing them to change their shape. For another example, until recently, RNA was not believed to be a direct player in cell control mechanisms, but now it is known that double-stranded RNA's (dsRNA's) can act, through the RNA interference (RNAi) effect, to disrupt ("turn-off" or "silence") genes. However, the central dogma remains the organizing principle: as usual in biology, the only general "theorem" is that every general fact has exceptions!

The DNA molecule is a double-stranded helix made of a sugar-phosphate backbone and nucleotide bases (Figure 3). Each strand carries the same information, which is encoded in the 4-letter alphabet $\{A, T, C, G\}$

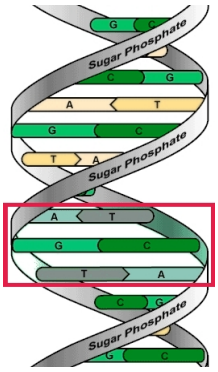

Figure 3: DNA; a codon shown in box

(the nucleotides Adenine, Thymine, Cytosine, and Guanine), in a "complementary" form (A in one strand corresponds to T in the other, and C to G). The two strands are held together by hydrogen bonds between the bases, which gives stability but can be broken-up for replication or transcription. One describes the letters in DNA by a linear sequence such as:

`gcacgagtaaacatgcacttcccaggccacagcagcaagaaggaggaatc`...

and genes (instructions that code for proteins) are substrings of the complete DNA sequence. (Besides genes, there are regulatory and start/stop regions that help delimit genes as well as determine if and when they should be "active". In addition, there are also regions that have other roles, such as coding for RNA that may not lead to proteins.) Because of its double-stranded nature, DNA is chemically stable, and serves as a good depository of information. One might think of DNA storage as a "hard disk" in a vague computing analogy.

2.2.2 RNA

The "read-out" of genetic information —bringing-in the instructions into working memory for execution, in our computer analogy— begins when DNA information is transcribed letter by letter into "RNA language." *Ribonucleic acid (RNA)* is a nucleic acid very similar to DNA, but less stable than DNA, and almost exclusively found in single-stranded form (with exceptions such as the RNA in some viruses). RNA language is basically the same as DNA's, with the minor (for us) detail that in RNA, the amino acid thymine is replaced with uracil, symbolized by the letter U. This process is known as *transcription*. The "copying-machine" is called *RNA polymerase*. A polymerase is, generally speaking, an *enzyme* —a type of protein that acts as a catalyst— that helps in the synthesis of nucleic acids. RNA polymerase is, thus, a polymerase that helps make RNA, more precisely *messenger RNA (mRNA)*.[2] A *promoter region* is a part of the DNA sequence of a chromosome that is recognized by RNA polymerase. In prokaryotes, the promoter region consists of two short sequences placed respectively 35 and 10 nucleotides before the start of the gene. Eukaryotes require a far more sophisticated transcriptional control mechanism, because different genes may be only active in particular cells or tissues at particular times in an organism's life; promoters act in concert with enhancers, silencers, and other regulatory elements.

[2]This description is over-simplified: in eukaryotic cells, an intermediate form of RNA called heterogeneous nuclear RNA (hnRNA) is produced first; then a process of "editing" gets rid of "introns" which are not part of the code for the desired protein, leaving the "exons" that are joined together to produce the actual mRNA, perhaps after insertion of some additional nucleotides.

2.2.3 Proteins

Proteins are the primary components of living things. Among other roles, they form receptors that endow the cell with sensing capabilities, actuators that make muscles move (myosin, actin), detectors for the immune response, enzymes that catalyze chemical reactions, and switches that turn genes on or off. They also provide structural support, and help in the transport of smaller molecules, as well as in directing the breakdown and reassembly of other cellular elements such as lipids and sugars. Ultimately, one might say that cell life is about proteins and how and when they are produced.

After transcription, *translation* is the next step in the process of protein synthesis and it is performed at the *ribosomes*. The information in the mRNA is read, and proteins are assembled out of amino acids (with the help of *transfer RNA (tRNA)*, which help bring in the specific amino acids required for each position). RNA language is translated into protein language by a mapping from strings written in the RNA alphabet $\Sigma_{\mathrm{n}} = \{U, A, G, C\}$ into strings written in the amino acid alphabet:

$$\Sigma_{\mathrm{a}} = \{A, R, D, N, C, E, Q, G, H, I, L, K, M, F, P, S, T, W, Y, V\}.$$

Every sequence of three letters in the RNA alphabet Σ_{n} is replaced by a single letter in the alphabet Σ_{a}. The genetic code explains how triplets (or *codons*, one of which is shown in Figure 3) of bases map into individual amino acids. The code, including full names and three and one-letter abbreviations, is shown in Figure 4. For

Alanine Ala A	GCU, GCC, GCA, GCG	Leucine Leu L	UUA, UUG, CUU, CUC, CUA, CUG
Arginine Arg R	CGU, CGC, CGA, CGG, AGA, AGG	Lysine Lys K	AAA, AAG
Asparagine Asn N	AAU, AAC	Methionine Met M	AUG
Aspartic Acid Asp D	GAU, GAC	Phenylalanine Phe F	UUU, UUC
Cysteine Cys C	UGU, UGC	Proline Pro P	CCU, CCC, CCA, CCG
Glutamine Gln Q	CAA, CAG	Serine Ser S	UCU, UCC, UCA, UCG, AGU, AGC
Glutamic Acid Glu E	GAA, GAG	Threonine Thr T	ACU, ACC, ACA, ACG
Glycine Gly G	GGU, GGC, GGA, GGG	Tryptophan Trp W	UGG
Histidine His H	CAU, CAC	Tyrosine Tyr Y	UAU, UAC
Isoleucine Ile I	AUU, AUC, AUA	Valine Val V	GUU, GUC, GUA, GUG
START	AUG, GUG	STOP	UAG, UGA, UAA

Figure 4: Genetic code

example, the codon AUG translates into M (Methionine). Thus, the DNA string $TACTCATTGCGC$ would first get transcribed into the RNA string $AUGAGUAACGCG$ (note the complementation, and replacing T by U), and would be then translated into the sequence $MSNA$ (Methionine-Serine-Asparagine-Alanine) of amino acids. The string AUG codes for the amino acid Methionine but also serves as a "start" codon: the first AUG in an mRNA indicates where translation should begin.

The *shape* of a protein is what largely determines its function, because proteins interact with each other, and with DNA and metabolites, through lego-like fitting of parts in lock and key fashion, transfer of small molecules, or enzymatic activation. Therefore, the elucidation of the three-dimensional *structure* of proteins is a central goal in biochemical research; this subject is studied in the fields of *proteomics* and *structural biology*. The *Protein Data Bank* (http://www.rcsb.org/index.html) based at Rutgers University, USA, serves as an online catalog of protein structures. Sometimes, protein structure can be gleaned through physical methods, such as X-ray crystallography or NMR spectroscopy. Very often, however, the structure of a protein P can only be estimated, based upon a comparison with an *homologous* protein Q whose structure has been already determined (as chemists say, "solved"). One says that P and Q are homologous if they are, in an appropriate

sense, close in amino acid sequence, or equivalently, in the DNA sequences for the genes coding for P and Q. One measure of closeness is Hamming distance (by how many "letters" do P and Q differ?), but more sophisticated measures used in practice include allowance for deletions and insertions of letters in P and Q. The rationale behind homology-based protein shape determination is that homologous proteins probably share a common evolutionary or developmental ancestry, and hence perform similar functions. Mathematical methods of computational biology (bioinformatics) play a central role in homology approaches; the *critical assessment of structure prediction methods (CASP)* competition compares methods from different researchers. Yet another set of techniques for elucidating the shape of proteins from their description as a linear sequence of amino acids is that of *energy minimization methods*. One views the protein-folding process as a gradient dynamical system, of which steady states are the stable configurations. This method is very difficult to apply, because of the complexity of the energy function, but has been useful for comparatively small proteins.

After translation, proteins are typically subjected to *post-translational modifications*, such as the addition of phosphate or methyl groups, or, in eukaryotic cells, *ubiquitination*, the process by which a protein is inactivated by attaching ubiquitin to it. Ubiquitin is a protein whose function is to mark other proteins for *proteolysis* (degradation), a process which occurs at the *proteasome*.

One of the key properties of proteins is that their shape (conformation) can be modified in a predictable fashion, as the consequence of interactions with other molecules. One often says that the protein has been "activated" as a result of such an interaction. For instance, Figure 5 shows, in schematic form, two conforma-

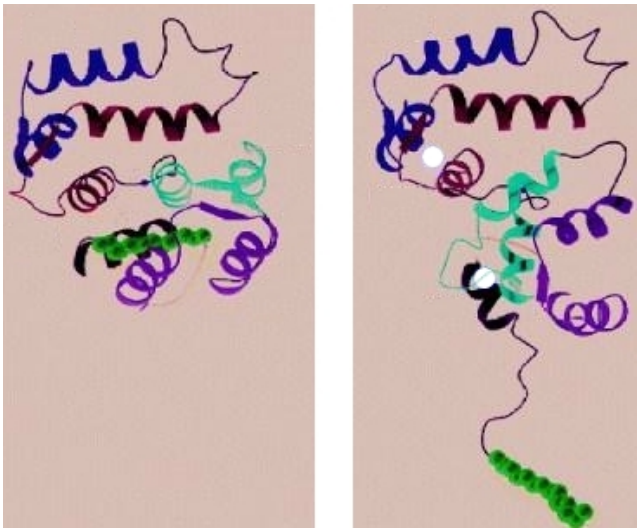

Figure 5: A protein in two conformations. Left one is Ca^{2+}-free. Right one is Ca^{2+}-bound

tions of the recoverin protein, the second of which comes about when two calcium ions have been inserted at appropriate places (white balls). Notice how the insertion of these ions makes an "arm" swing out. Depending on the position (extended or not) of this arm, different interactions of this protein with other players in the cell will occur.

2.3 Proteins act as Sensors, Signal Relayers, and Actuators

Conformation changes in proteins typically happen in response to intracellular or extracellular ligand binding events, or because of binding with other proteins. (To *bind* means to reversibly join; *ligands* are small molecules that bind with larger molecules, typically proteins.) Two noteworthy instances of activation are provided by receptors and by phosphorylation reactions.

Receptors are proteins that act as the cell's sensors of outside conditions, relaying information to the inside of the cell. A receptor is typically made up of three parts. The *extracellular domain* ("domains" are parts of a protein) is exposed to the exterior of the cell. Extracellular ligands, such as growth factors and hormones, bind to receptors, most of which are designed to recognize a specific type of ligand. The *transmembrane domain* serves to "anchor" the receptor to the membrane. Finally, a *cytoplasmic domain* helps initiate reactions inside the cell in response to exterior signals, by interacting with other proteins. There is a special class of receptors which constitute a common target of pharmaceutical drugs: *G-protein-coupled receptors (GPCR's)* (Figure 6). The name of these receptors arises from the fact that, when their conformation changes in response to a ligand

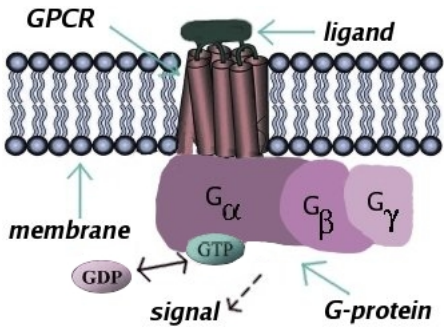

Figure 6: G-protein-coupled receptor and G-protein

binding event, they activate G-proteins, so called because they employ *guanine triphosphate* and *diphosphate (GTP* and *GDP)* in their activity. GPCR's are made up of several subunits (G_α, G_β, G_γ) and are involved in the detection of metabolites, odorants, hormones, neurotransmitters, and even light (rhodopsin, a visual pigment).

Another example of activation is *phosphorylation*. *Adenosine triphosphate (ATP)* is a nucleotide that is the major energy currency of the cell. An *enzyme* is a protein that catalyzes a chemical reaction. Phosphorylation is a chemical reaction in which an enzyme X —called a *kinase* when playing this role— transfers a phosphate group (PO_4) from a "donor" molecule such as ATP to another protein Y, which becomes "activated" in the sense that its energy is increased. Once activated, protein Y may then influence other cellular components, including other proteins, itself acting as a kinase, or it may take an appropriate shape that allows it to to bind with yet another protein or to a segment of DNA so as to initiate, enhance, or repress expression of a gene. Normally, proteins do not stay activated forever; another type of enzyme, called a *phosphatase*, eventually takes away the phosphate group; see Figure 7. In this manner, signaling is "turned off" after a while, so that the system is

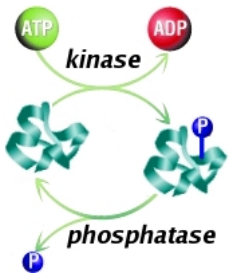

Figure 7: Phosphorylation and de-phosphorylation

ready to detect new signals.

Receptors and enzymatic cascades act in concert. Binding of extracellular ligands triggers signaling through a series of chemical reactions inside the cell, carried out by enzymes and often relayed by smaller molecules called *second messengers*. In this manner, regulatory pathways can be either turned "on" and "off" or mod-

ulated, and transcription of particular sets of genes may be started and stopped in response to environmental conditions. Figure 8 ([19]) illustrates one such pathway, which involves GPCR activation as well as signaling through a MAPK cascade (more on MAPK cascades below).

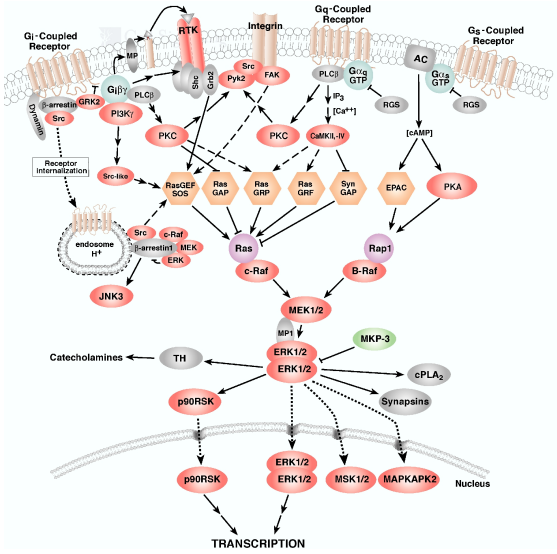

Figure 8: A GPCR pathway

The animation at *http://biocreations.com/pages/mapk.html* is strongly recommended as an illustration of signaling pathways.[3]

As another illustration, consider the diagram shown in Figure 9, extracted from the paper [46] on cancer research, describing the top-level schematics of a wiring diagram of signaling circuitry in the mammalian cell. The illustration shows the main signaling pathways for growth, differentiation, and apoptosis (commands which instruct the cell to die). Highlighted in red are some of the genes known to be functionally altered in cancer cells. Of course, such a figure, compared for example with the more detailed biochemical pathway shown in Figure 8, leaves out a lot of information, some known but omitted for simplicity, and some unknown. Much of the system has not yet been identified, and the functional forms of the interactions, much less parameters, are only very approximately known. However, data of this type are being collected at an amazing rate, and better and better models are being obtained constantly.

Both of the above examples were from eukaryotes. We now turn to one from a prokaryote. *Chemotaxis* is the term used to describe movement, in bacteria as well as other organisms, in response to chemoattractants or repellants, such as nutrients and poisons, respectively. *E. coli* bacteria (Figure 10) are single-celled organisms, about 2 μm long, which possess up to six flagella for movement. Chemotaxis in *E. coli* has been studied extensively. These bacteria can move in basically two modes: a "tumble" mode in which flagella turn clockwise and reorientation occurs (Figure 11, left), or a "run" mode in which flagella turn counterclockwise, forming a bundle which helps propel them forward (Figure 11, right). The motors actuating the flagella are made up of several proteins. In the terms used by Berg in [15], they constitute "a nanotechnologist's dream," consisting as they do of "engines, propellers, ..., particle counters, rate meters, [and] gear boxes." Figure 12 shows an actual electron micrograph and a schematic diagram of a flagellar motor. The signaling pathways involved in *E. coli* chemotaxis are fairly well understood. Aspartate or other nutrients bind to receptors, reducing the rate at which a protein called CheA ("Che" for "chemotaxis") phosphorylates another protein called CheY transforming it into CheY-P. A third protein, called CheZ, continuously reverses this phosphorylation; thus, when ligand is present, there is less CheY-P and more CheY. Normally, CheY-P binds to the base of the motor,

[3] signaling cascade animation played at this point of lecture

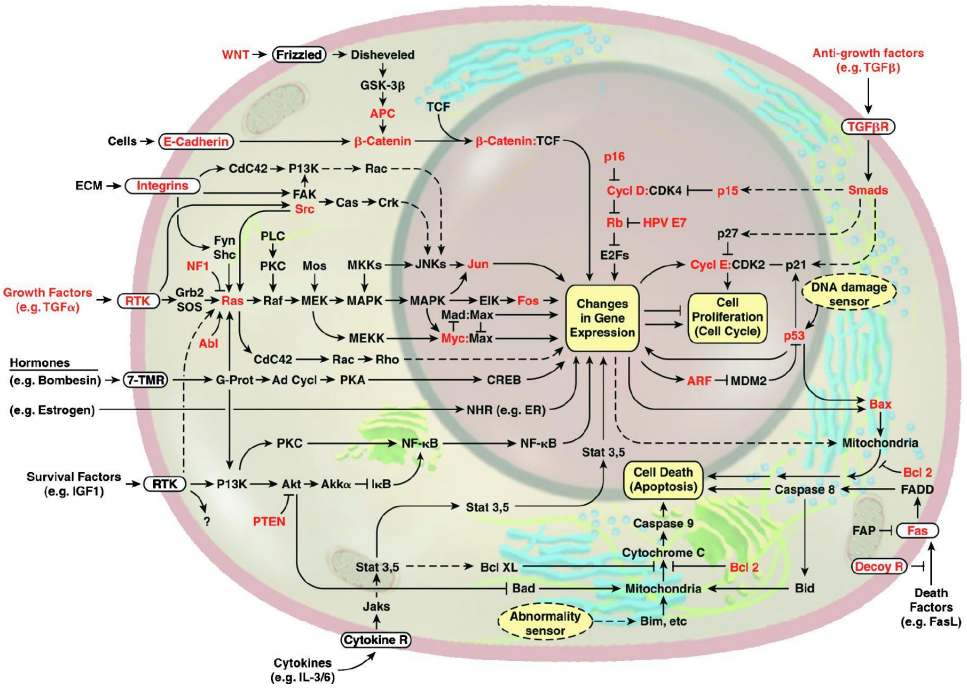

Figure 9: Signaling circuitry of the mammalian cell from [46], reprinted with permission from Elsevier

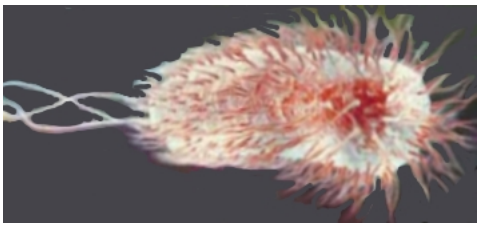

Figure 10: *E. coli* bacterium

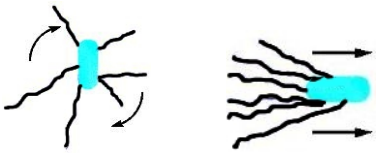

Figure 11: *E. coli* tumbling: flaggela apart. Running: flaggela in bundle

helping clockwise movement and hence tumbling, so the lower concentration of CheY-P has the effect of less tumbling and more running (presumably, in the direction of the nutrient). A separate feedback loop, which includes two other proteins, CheR and CheB, causes adaptation to constant nutrient concentrations, resulting in a resumption of tumbling and consequent re-orientation. In this manner, *E. coli* performs a stochastic gradient search in a nutrient-potential landscape. Figure 13 shows a schematic diagram of the system responsible for chemotaxis in *E. coli*.

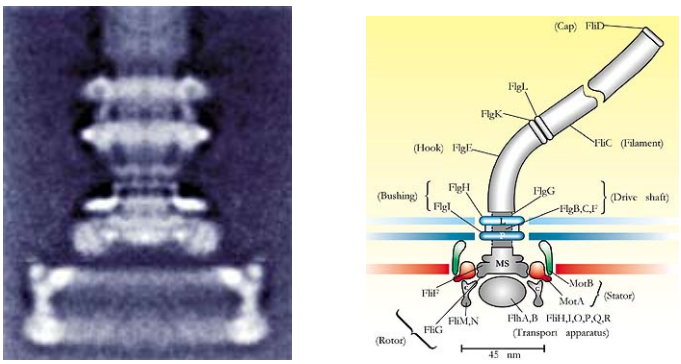

Figure 12: Electron micrograph and diagram of flagellar motor, reprinted with permission from [15]

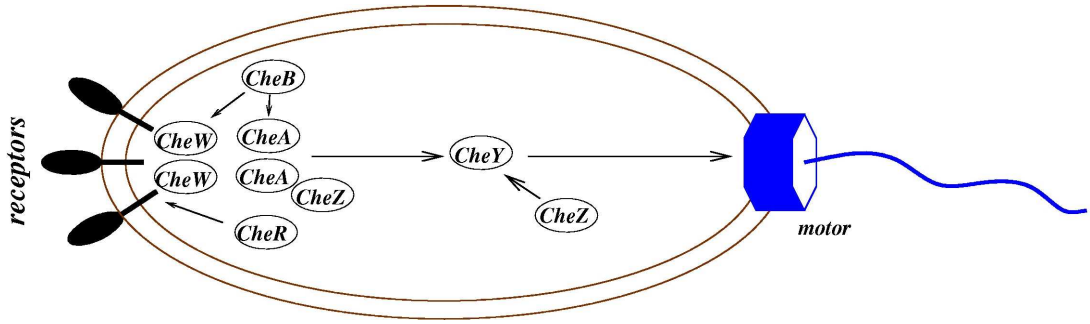

Figure 13: *E. coli* chemotactic circuit

2.4 Measurement Techniques

Massive amounts of data are being generated by genomics and proteomics projects, thanks to sophisticated genetic engineering tools (gene knock-outs and insertions, PCR) and measurement technologies (fluorescent proteins, microarrays, blotting, FRET). *Polymerase chain reaction (PCR)* is a technique that amplifies DNA (typically a gene or part of a gene). Creating multiple copies of a piece of DNA, which would otherwise be present in too small a quantity to detect, PCR enables the use of measurement techniques. Let us briefly discuss a couple of these measurement technologies, in order to provide an idea of their power as well as their severe limitations.

Suppose that we wish to know at what rate a certain gene X is being transcribed under a particular set of conditions in which the cell finds itself. Fluorescent proteins may be used for that purpose. 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 gene*. The GFP gene is inserted (cloned) into the chromosome, adjacent to or very close to the location of gene X, so both are controlled by the same promoter region. Thus, gene X and GFP are transcribed simultaneously and then translated (Figure 14), so by measuring the intensity of the GFP light emitted one can estimate how much of X is being expressed.

Fluorescent protein methods are particularly useful when combined with *flow cytometry*. Flow Cytometry devices can be used to sort individual cells into different groups, on the basis of characteristics such as cell size, shape, or amount of measured fluorescence, and at rates of up to thousands of cells per second. In this manner, it is possible, for instance, to count how many cells in a population express a particular gene under a specific set of conditions.

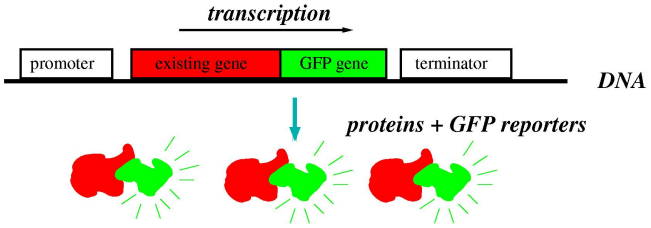

Figure 14: GFP

A set of technologies collectively referred to as *gene arrays* (DNA chips, DNA microarrays, Affymatrix gene chips) provide high-throughput methods for simultaneously monitoring the activity levels of thousands of genes, thus providing a snapshot of the current gene expression activity of a cell (Figure 15). An array

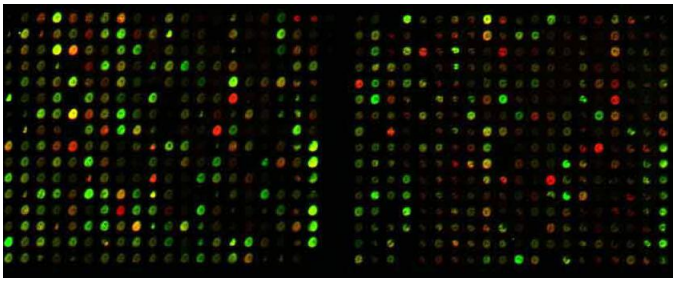

Figure 15: Gene array

is built using robotics and imaging equipment, very much as in electronic chip fabrication. The array has in each location (i, j) a detector "tuned" to a particular gene or small sequence of nucleotides X_{ij}. This detector (the usual name is a "target") is the complement $\overline{X}_{ij}$ of X_{ij} or, more likely, of a subsequence of X_{ij}. (More precisely, one wants to find out how much of a specific X's mRNA is being transcribed. The first step is to reverse-transcribe RNA to DNA, to obtain called complementary DNA (cDNA), and then PCR-amplify it. We omit details here, since we only want to explain the basic principle.) Because of hybridization, that is, the A-T and G-C base pairings for DNA, X_{ij} should "stick" to its complement $\overline{X}_{ij}$. This allows one to estimate the presence and abundance of each X_{ij} in a sample. In order to be able to read the information in the different array positions, the sequences X_{ij} being tested for are first radioactively or fluorescently tagged, so that one can simply measure how much has accumulated at each position i, j. Pattern recognition, machine learning, and control-theory tools such as clustering, Bayesian networks, and identification theory —especially when time-dependent data is available— can be and are used infer information about dynamic interactions among genes, and to sort out which particular sets of genes are triggered simultaneously or in a sequence (*co-expression* analysis) in response to different environmental factors or disease states. In control-theory language, we might think of gene arrays as giving a vector-valued output, in contrast to a technology such as GFP which provides merely a scalar value.

Actually, it is difficult to obtain absolute measurements with gene arrays, due to uncertainties in the PCR and hybridization processes. Rather, the method is often used in a comparative fashion. Gene array experiments can be done for different cell types in the same organism, for the same cell types under different experimental conditions, or even for comparing cells from two organisms, perhaps one of them having an engineered mutation of the original one. A fascinating application is the comparison of abnormal (e.g., cancerous) and normal cells, obtaining in that manner a gene expression "signature" that might be used for diagnosis.

A *Western blot* allows one to detect the presence of a specific protein, or a small number of them, in a sample

taken from an experiment.[4] The proteins extracted from the sample, together with a small number of antibodies which recognize only specific proteins, are placed on membranes and allowed to interact. Different methods, for instance radioactive labeling of stains, are then used in order to visualize the results. As an example, Figure 16, taken from [73], shows Western blot data from an experiment in which three proteins (Cdc25, Wee1, and MAPK) have been observed under different conditions (concentrations 0, 25nM, etc.) of another protein named Δ65-cyclin B1, during two experiments (labeled "going up" and "coming down" in the figure). The higher placements on the blot correspond in this case to the relative abundance of the phosphorylated form of the protein; for example, phosphorylated Cdc25 is more abundant in the "100" than in the "0" lanes.

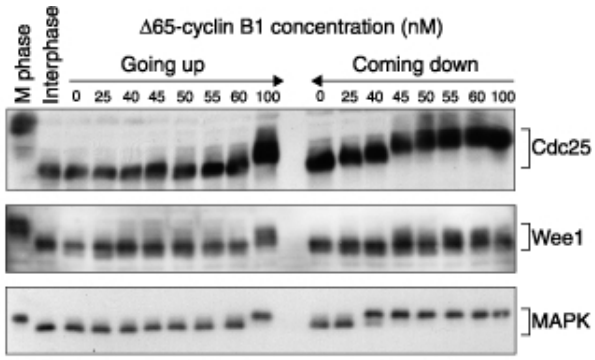

Figure 16: Western blots

2.5 Limitations

Notwithstanding the power of the techniques just described, GFP, arrays, and blots, they are intrinsically noisy, because of chemical interactions in blots, production errors in arrays, or other sources of interference. In addition, the resulting measurements have low precision: very few bits of information can be extracted from data such as that shown in Figures 15 or 16. These limitations of imprecision and noise are sometimes ignored in systems biology modeling, but it is obviously pointless to try to tightly fit model parameters to such data. On the other hand, for certain types of quantities, such as the amount of calcium in a cell, currents through channels, or certain enzyme concentrations, there are other techniques that may result in higher precision measurements. In such cases, parameter fitting is more reasonable.

The field suffers from what has been called a *data-rich/data-poor* paradox: while on the one hand a huge amount of *qualitative* network (schematic modeling) knowledge is available, as evidenced by figures such as 8 and 9, on the other hand little of this knowledge is *quantitative*, at least at the level of precision demanded by most control theoretic tools of analysis. The problem of exploiting this qualitative knowledge, and effectively integrating relatively sparse quantitative data, is among the most challenging issues confronting systems biology.

2.6 Model Organisms

Since many organisms follow the same basic principles, biologists have concentrated on a small number of *model systems*. This allows them to focus on specific systems, easing comparisons and facilitating sharing of research results. Different aspects may be easier to study in different model organisms (embryonic cycles in

[4] "Southern" blots are techniques for detecting DNA, and "Northern" blots for detecting RNA. The names originated with the first of these, which was developed by a UK biologist named Southern.

frog eggs, differentiation and development in flies, aging in worms), by taking advantage of fast breeding or speed of maturation.

As cataloged in the US National Institutes of Health website (http://www.nih.gov/science/models), the main mammalian models are the mouse and rat, and the main non-mammalian models are *S. cerevisiae* (budding yeast), *Neurospora* (filamentous fungus), *D. discoideum* (social amoebae), *C. elegans* (round worm), *D. melanogaster* (fruit fly), *D. rerio* (zebrafish), and *Xenopus* (frog). In addition, a popular plant model is *Arabidopsis* (a small flowering plant, member of the mustard family).

Most mathematical modeling, signal processing, and feedback control studies have been done specifically for one or another of these model systems.

3 Cells as Dynamical Systems

The term *genotype* refers to the genetic blueprint encoded in the DNA of a given individual, while *phenotype* refers to the actual observable physical manifestations of that information. A *single nucleotide polymorphism (SNP)*, that is, a change (mutation) in a single letter in an individual's DNA, may not have a phenotypical consequence, or it might have a catastrophic one, as is the case with cystic fibrosis in humans. Moreover, distinct species may be relatively close in genotype, yet be very far in other characteristics; for example, humans and chimpanzees are close to 99% genetically identical. Thus, differences in genotype can be tremendously amplified into phenotype. But, even accounting for environmental factors ("inputs" to the system), this amplification would seem to be somewhat inconsistent with the Central Dogma. After all, the mapping "genome $\mapsto$ proteome" is quite "continuous" in an intuitive sense and proteins determine the organism. One might then ask how large discontinuities arise.

One major contributing factor is that a cell behaves as a *nonlinear dynamical system*. As we discussed, proteins interact among themselves, both directly, through enzymatic action or through binding, as well as indirectly, through their control of gene expression. Each of these modes of interaction may involve *feedback loops*. Feedback is properly understood as a dynamic phenomenon, where quantities, such as concentrations of proteins, RNA, metabolites, and other cell substances are seen as functions of time.

Feedback in gene expression, to take one example, is critical to the cell's function ([29, 105]). A *transcription factor* is a protein that directs when –and possibly how many times– a gene is to be transcribed, by binding to DNA at a specific promoter or other regulatory region. Thus, a protein A may inhibit or enhance transcription of the RNA that codes for some other protein B, while B may in turn influence the production of A. Combinations of such influences are possible, as illustrated in Figure 17, in which proteins A and B must both be present in order for gene C to be active (an "and" gate in Boolean terms); the boxes labeled $P_A, P_B, P_{C_1}, P_{C_2}$ indicate

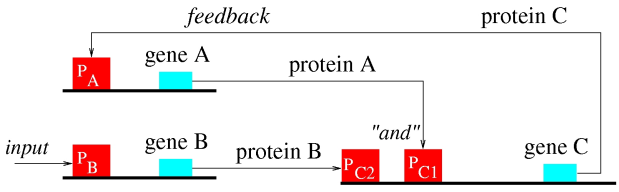

Figure 17: Proteins feed back into gene expression

regulatory sites.

The various modes of interaction are closely related: an enzymatic signal transduction network may direct the activation of a transcription factor, or a reaction of protein *dimerization* —binding of two proteins to each other— may be required for transcription factor activation. For example, the diagram in Figure 18 shows a

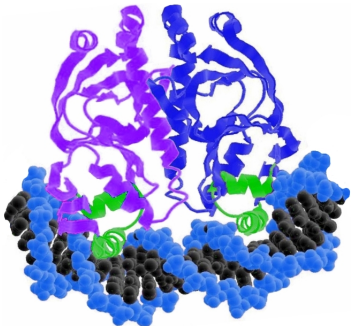

Figure 18: A homodimer, bound to DNA

homodimer —that is, a dimer consisting of two proteins of the same type, in this case catabolite gene activator protein (CAP), which is one of over 300 transcription factors in *E. coli*— bound to DNA, which in turn helps RNA polymerase bind and initiate transcription.

Before continuing, let us very briefly digress to mention two important sources of nonlinearities. One of them is dimerization. If a dimer consisting of a molecule of P and a molecule of Q plays a role in a reaction, then we must keep track of the amount of the dimer, let us call it D. Now, a molecule of D forms whenever a molecule of P interacts with a molecule of Q. Assuming that the medium is well-mixed —for instance due to Brownian motion— the probability that such an interaction will occur is proportional to the product of the concentrations of P and of Q. Thus, in a differential equation model that keeps track of concentrations, a product term $p(t)q(t)$ will be required in order to represent this dimerization. In particular, for homodimers one may expect to see a term like $p(t)^2$. In general, one calls an exponent appearing in this fashion a *cooperativity* index. Even higher order monomials may appear; for example, it is known that receptors in *E. coli* tend to aggregate in large numbers. Another important way in which nonlinearities appear is through saturation effects, for instance if an enzyme E catalyzes the conversion of a substrate S into a product P and the enzyme is in short supply, there will be a maximal speed at which the reaction can take place.

3.0.1 Bifurcations

We wish to argue that dynamical phenomena are a main contributing factor to the appearance of discontinuities. Mathematically, such discontinuities are described as bifurcations, where a small change in a parameter results in completely different steady state behavior. (Another possibility, probably less important in this context, is the existence of chaotic dynamics, which exhibit sensitive dependence to initial conditions, so that small differences in initial states result in quickly diverging trajectories, even on finite time intervals.) Let us give a simplified illustration of the biochemical role of bifurcations; much more complicated, but totally analogous, mathematical models appear, for example, in papers dealing with embryonic development or signaling pathways. Suppose that $p(t)$ denotes the dimensionless concentration ($0 \leq p \leq 1$), at time t, of the protein product P of some gene, whose presence results in some observable characteristic of the individual, and that p evolves in time according to the following differential equation:

$$\frac{dp}{dt} \;=\; p^2(1-p) - kp\,.$$

The negative term corresponds to degradation, and the first term to formation by an autocatalytic process, with the square term representing a dimerization. The parameter k represents the activity of some enzyme that facilitates the degradation of p. Let's assume that $p(0) = 0.5$, an initial condition that might have been set up by another process. (In embryonic development, some of the initial conditions are set by chemical gradients placed by the mother on the fertilized egg, cf. [103].) If $k > 1/4$, then $f(p) = p(-p^2 + p - k)$ is always negative, so $p(t) \to 0$ as $t \to \infty$, that is, complete degradation of p results. On the other hand, if $k < 1/4$,

then $f(p) = p(-p^2 + p - k)$ has two roots $p_- < 0.5 < p_+$, so $p(t) \to p_+$. Thus *a slight perturbation of the parameter k will have a drastic, discontinuous, effect on the phenotype.*

3.0.2 Activation and Inhibition

It is common to classify biochemical interactions as negative (inhibitory) or positive (activating). Suppose that we consider two interacting chemicals P and Q. The rate of change of P may be affected by the concentration of Q in several different ways. For example, Q might be an enzyme that helps catalyze the production of P, or a protein whose presence triggers the expression of the gene that produces P; in this case, we say that the effect of Q on P is *positive* on P, or that Q is an *activator* of P. Alternatively, Q might be an enzyme that helps degrade P, or a protein that represses the gene that produces P, in which case we say, instead, that Q has a *negative* effect on P, or that Q *inhibits* P. Similarly, one can define activation or inhibition of Q by P. Of course, it could happen that the effect of P on Q (or of Q on P) is ambiguous, and depends on the actual concentrations of P and Q, or even of other species. However, it is often —though certainly not always— the case that biochemical models are *sign-definite*, by which we mean that this change of sign cannot happen: either P always inhibits Q or P always activates Q. As an example, take Figure 9, part of which we provide a closer look of in Figure 19. In

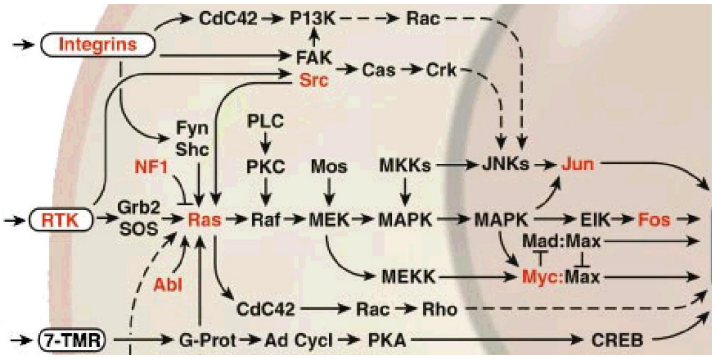

Figure 19: Zooming-in on Figure 9

this picture, the arrows "$\rightarrow$" indicate activation, and the symbols "$\dashv$" indicate inhibition.

To give precise definitions, one needs to settle upon a type of model for the concentrations $p(t)$ and $q(t)$ as functions of time t: ordinary or partial differential, probabilistic, Boolean, or hybrid equations. For concreteness, suppose that the pair of ordinary differential equations

$$\begin{aligned} \dot{p} &= f(p,q) \\ \dot{q} &= g(p,q) \end{aligned}$$

adequately describes the interaction between P and Q (as usual in control theory, using dot for time derivatives and omitting "t" arguments). Then, Q is an activator of P if the partial derivative $\frac{\partial f}{\partial q}(p,q)$ is positive, or at least nonnegative, everywhere, and Q is an inhibitor of P if this derivative is negative, or at least nonpositive, everywhere. The non-sign-definite case would be that in which the partial derivative is positive for some values of the state variables (p,q) and negative for others, as with the equation $\dot{p} = (1-p)q$, where $\frac{\partial f}{\partial q} = 1 - p$ is positive if $p < 1$ and negative if $p > 1$.

3.0.3 Two-Species Interactions

As the number of proteins and other species increase, the complexity of feedback loops and dynamics exhibited by biochemical networks can be, in principle, quite arbitrary. However, some of the main behaviors in

which biologists have focused their interest arise already in systems that involve just two interacting chemicals, Figure 20.

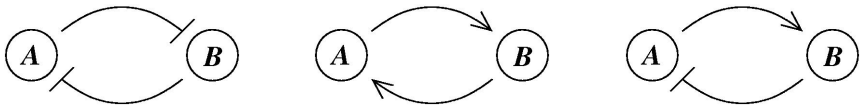

Figure 20: Mutual inhibition. Mutual activation. Activation-inhibition

In *mutual inhibition*, each species inhibits the other one. If some external input signal helps to transiently increase the concentration of A sufficiently over that of B, then A will repress B. Since B is at a low concentration, it will not repress A. Assuming that A can maintain its high level —due for example to some autocatalytic reaction, or to the influence of other variables not shown— this situation will persist until such a time when some other external factor allows B to gain an upper hand over A. The system will, therefore, memorize which of the two components, A or B, was last activated externally; this "toggle-switch," analogous to similar ones found in electronic devices, plays a central role in differentiation and other biological forms of memory. See for instance older work on the lambda phage lysis-lysogeny switch and the hysteretic *lac* repressor system [71, 74], as well as more recent references such as [40, 18, 80, 14].

In *mutual activation*, each species activates the other. Now, if some external input signal helps to transiently increase the concentration of A, then B will be activated by A, and B will, in turn, enhance A even more. In effect, a sufficiently large external signal, applied to either A or B, results in a large increase in both A and B. (If the signal is not strong enough, we will assume that A and B stay small.) This mode of positive feedback appears in biomolecular systems that amplify signals, as well as systems that produce a "binary" response to external stimuli, and it is thought to play a role in cell decision-making.

Finally, a net negative feedback as in *activation/inhibition* loops is, as usual in control theory, the mechanism responsible for set-point regulation, or as biologists say, *homeostasis*. It plays a role also in turning signals "off" after activation: many cell signals are too expensive metabolically to be maintained at a high level.

These behaviors are associated with different phase-space pictures, which we discuss now, for concreteness, for ordinary differential equation models.

3.1 Phase Spaces and Step Response

Three of the main types of phase-space behaviors that have attracted particular attention from biologists studying biomolecular dynamics are: systems with a unique stable state, systems with multiple attracting states, and limit cycle oscillators, cf. Figure 21. These three types of behaviors are intimately linked, and often give rise to

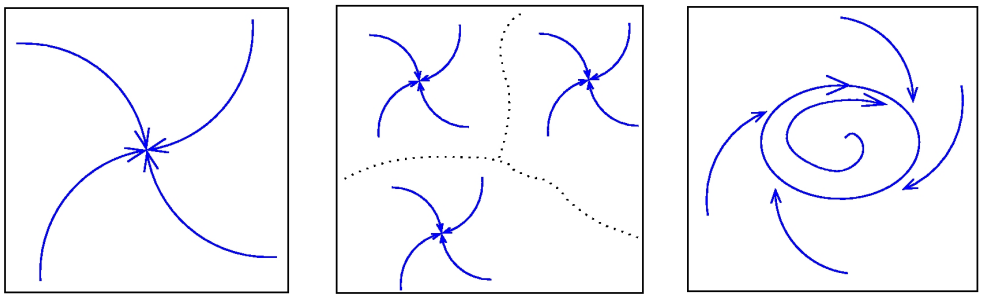

Figure 21: One or multiple steady states; Limit cycles

each other, as we will discuss.

Uniqueness of steady states, and globally asymptotic stability, are quite common among simple biochemical reactions, although it is not always easy to prove theorems insuring this behavior (we discuss some such results later). Systems with multiple attractors arise in many forms, a typical one of which is the interaction between two processes, such as formation and degradation, each of which by itself would lead to global stability. Relaxation, or hysteresis-driven, oscillators are those in which to a system with multiple attractors one adds a slow parameter adaptation law. Other oscillators arise through a Hopf bifurcation phenomenon –basically an unstable linear oscillator, plus a nonlinear term that prevents escape to infinity and thus confines trajectories– from negative feedback loops around otherwise mono-stable systems.

The transitions (bifurcations) between qualitative behaviors such as mono- and multiple-stability, or the onset of oscillations, are phenomena which frequently arise when parameters in systems are modified. In molecular biology modeling, a parameter may typically represent a concentration of an external ligand, a voltage applied to a voltage-gated channel, the concentration of a signaling molecule (as an input to a cellular subsystem), an enzyme concentration affecting a reaction, or the degree of effective cooperativity (Hill coefficient) of a reaction.

For example, suppose that the rate of change of the concentration of some substance p has the following form:

$$\dot{p} \;=\; \frac{V_{max}u}{k_m + u} \;-\; kp$$

where we fix the parameters V_{max}, k_m, and k, and where u is a parameter, not fixed yet, which might correspond, for instance, to the concentration of substrate that is used in making p. The term $-kp$ is a degradation term, while the first term is a *Michaelis-Menten* formation term. (Michaelis-Menten kinetics are really a singular perturbation reduction of a more complicated underlying enzymatic reaction, see e.g. [34] for details.) Note that V_{max} is the maximum possible speed of the formation reaction, while k_m ("m" for middle) is the concentration of u for which the rate happens to be $V_{max}/2$. Now assume that u is fixed at some value u_0. The concentration $p(t)$ will then, from any initial condition $p(0)$, converge to the steady state $p_0 = \frac{(V_{max}/k)u_0}{k_m+u_0}$. In a typical set of experiments, a biologist or biochemist will set the concentration to a given value u_0, let the system relax to the corresponding steady state p_0, and repeat for various values of u_0, thus obtaining a plot of p_0 against different such u_0's. In Figure 22 we show the plot (with $V_{max} = 1$, $k_m = 0.25$, $k = 1$) for the above

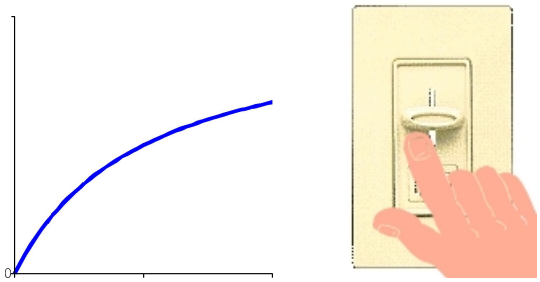

Figure 22: Hyperbolic steady-state response

example. We'll call this graph, using control-theory terminology, the *steady state response to step inputs*, where we think of u_0 as the magnitude of a constant input applied to the system. Depending on the context, this plot might be called a *dose-response curve* or *receptor activity plot* when u represents a concentration of ligand and p the level of some indicator of receptor activity, a *steady-state phosphorylation level* plot when u represents a signal that affects the phosphorylation level of a protein, and so forth. The response in this example is *graded* in the sense that it is proportional to the parameter u_0, at least over a large range of values u_0, even though it eventually saturates. It is said to be a *hyperbolic* response, in contrast to a *sigmoidal* response as in Figure 23. A sigmoidal response arises typically from a reaction such as:

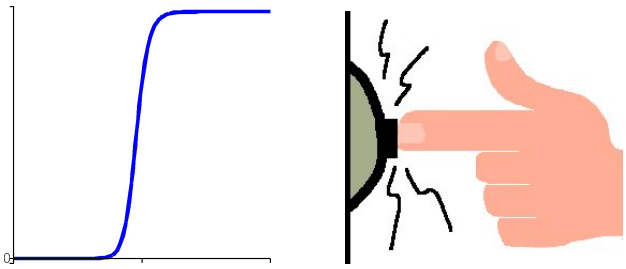

Figure 23: Sigmoidal steady-state response

$$\dot{p} \;=\; \frac{V_{max}u^r}{k_m^r + u^r} \;-\; kp$$

where the *Hill coefficient* r is greater than one (in our figure, we used $r = 20$, $V_{max} = 1$, $k_m^r = 0.4$, and $k = 1$). The parameter r is a cooperativity index. The sharp increase, and saturation, means that a value of u_0 which is under some threshold (roughly, $u < k_m$) will not result in an appreciable result ($p_0 \approx 0$, in steady state) while a value that is over this threshold will give an abrupt change in result ($p_0 \approx V_{max}/k$, in steady state). While the first example, when we think of u_0 as displacement of a slider or button, is analogous to the behavior of a light-dimmer, the second one is closer to that of a doorbell. (We do not define here precisely the difference between sigmoidal and hyperbolic responses. One possible definition is in terms of inflection points in the graph. But there is no need to be formal, since we want to keep the discussion intuitive at this point.)

Sigmoidal responses are characteristic of many signaling cascades, which display what biologists call an *ultrasensitive* response to inputs. If the purpose of a signaling pathway is to decide whether a gene should be transcribed or not, depending on some external signal sensed by a cell, for instance the concentration of a ligand as compared to some default value, such a binary response is required. Cascades of enzymatic reactions can be made to display ultrasensitive response, as long as at each step there is a Hill coefficient $r > 1$, since the derivative of a composition of functions $f_1 \circ f_2 \circ \ldots \circ f_k$ is, by the chain rule, a product of derivatives of the functions making up the composition ([41]). Thus, the slopes get multiplied, and a steeper nonlinearity is produced. In this manner, a high effective cooperativity index may in reality represent the result of composing several reactions, perhaps taking place at a faster time scale, each of which has only a mildly nonlinear behavior.

In practice, steady-state step response curves are interpolated from a number of measurements taken for various values of u_0. For a concrete, although relatively old, example, we show in Figure 24, taken from [33], a (log scale) plot of the degree of cAMP receptor modification after 15 minutes of constant exposure to the stimulant cAMP, in *Dictyostelium*. The locations of the black circles are obtained by reading the Western blots shown in the inset.

We mentioned that systems with multiple attractors sometimes arise through the interaction of formation and degradation processes. A typical way in which this happens is as follows. Suppose that the output y of a system, for example $y = p$ in the example that we have been considering, is fed-back into the input u, as shown diagrammatically in Figure 25(a). Physically, we are dealing with an autocatalytic process, and may think simply of u being equal to p (this could happen for example if p helps promote its own transcription) or perhaps there could be a more complicated positive feedback pathway from p to u. Mathematically, we substitute $u = y$ into $\dot{p} = \frac{V_{max}u^r}{k_m^r+u^r} - kp$ (where $r = 1$ or $r > 1$), and obtain the closed-loop equation:

$$\dot{p} \;=\; \frac{V_{max}p^r}{k_m^r + p^r} \;-\; kp\,.$$

We plot in Figure 26 both the first term (formation rate) and the second one (degradation), in cases where $r = 1$ (left) or $r > 1$ (right). Let us analyze the solutions of the differential equation. In the first case, $r = 1$, for small p the formation rate is larger than the degradation rate, while for large p the degradation rate exceeds

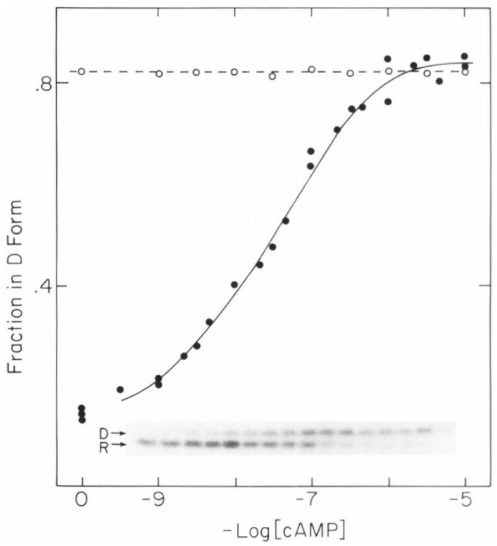

Figure 24: Example of steady-state response, from [33]

Figure 25: (a) Feed-back $u = y$; (b) Feed-back $u = g \cdot y$

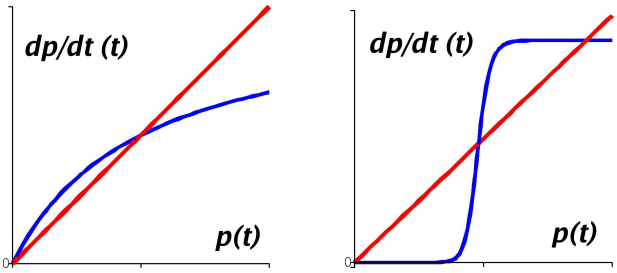

Figure 26: Bistability arises from sigmoidal formation rates

the formation rate; thus, the concentration $p(t)$ *converges to a unique intermediate value.* In the second case, however, the situation is more interesting: for small p the degradation rate is larger than the formation rate, so the concentration $p(t)$ *converges to a low value*, but, in contrast, for large p the formation rate is larger than the degradation rate, and so the concentration $p(t)$ *converges to a high value* instead. In summary, two stable states are created, one low and one high, by this interaction of formation and degradation, if one of the two terms is sigmoidal. (There is also an intermediate, unstable state.) *These facts are totally elementary, but they serve to motivate a theory based upon monotone systems, to be explained later, which provides a far-reaching generalization.*

Whether, under feedback, a mono-stable or a multi-stable system results, depends on the shape of the curves, which in turn is determined by the numerical values of the parameters. For example, the hyperbolic case, shown in the left panel of Figure 26, corresponds to $r = 1$, while $r \gg 1$ tends to produce pictures like the one shown in the right panel.

Other parameters also play a role. Let us consider a situation where the strength of the feedback can be modulated in some fashion, for example due to some additional transcriptional or enzymatic control. The simplest case (the theory works equally well in more complex scenarios) is when u is proportional to the output y: $u = g \cdot y$, where "g" (a "feedback gain" in engineering) is a parameter than quantifies the proportion (amplification, if $g > 1$) of y that is fed back as input, That is, instead of closing the loop simply with $u = y$ as in Figure 25(a), we now wish to study the effect of a more general feedback $u = g \cdot y$, where $g \neq 1$, as in Figure 25(b).

For example, consider the sigmoidal curve shown in the left panel of Figure 27, showing the steady-state

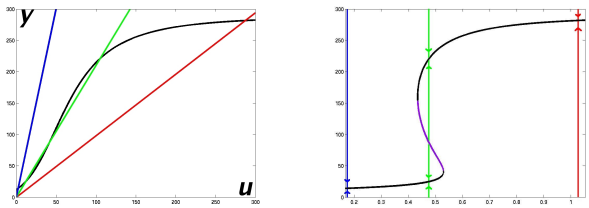

Figure 27: Open-loop step-response and three feedback gains; Corresponding bifurcation diagram

step response $y = k(u)$ of a certain system that we will study later. For any fixed value of g, the steady states of the closed-loop in Figure 25(b) are in one-to-one correspondence with those pairs (u, y) for which both $y = k(u)$ and $u = gy$, that is to say, the intersections of the graph of $y = k(u)$ with the line $y = (1/g)u$. In particular, we show in the left panel of Figure 27 the lines $y = (1/g)u$ with slopes corresponding to the three special values $g = 1/0.98$, $g = 1/2.1$, and $g = 1/6$. The middle line would correspond to a bistable case as in the right panel of Figure 26, while the other two lines correspond to cases where a single steady state will occur, either a "low y" or a "high y" one. We may plot the y-coordinates of these intersections against the values of the gains g. Observe that this plot, the bifurcation diagram, can be easily obtained by a projective transformation from the data given by the steady state response $y = k(u)$: it is simply given, in parametric form, as the set of pairs $\left(\frac{u}{k(u)}, k(u)\right)$ parametrized by possible input values u. See the right panel of Figure 27. Thus, if k is obtained from experimental data, the bifurcation diagram can be immediately derived from it.

An intuitive way of thinking of the dependence of the steady state on the parameter g is by viewing g as the force being applied on a light switch (let us say, positive means up, and negative means down) as in Figure 28. A strong enough positive force will turn the light on, and a strong enough negative force will turn it off, no

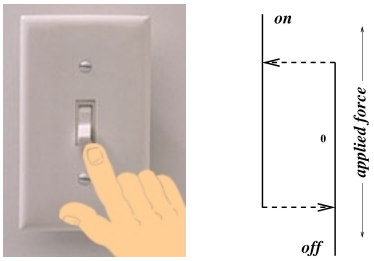

Figure 28: Hysteretic behavior

matter how we started. An intermediate value will have an effect that depends on the initial state: if the switch is only partially up, but the light is off, a small force will leave it off; if it is on, it will stay on. This is the bistable case, where the steady state attained depends on the initial state. *Hysteresis* is the term used to describe the phenomenon in which the actual steady state depends on the history of the system. One of the main roles of such hysteretic behavior is in producing oscillations. Imagine an indecisive individual, who, when the light is

off starts applying a higher and higher upward pressure on the switch, but, when the light turns on, changes his mind and starts applying a downward pressure, repeating the process forever. The resulting oscillation is called a hysteresis-based or relaxation oscillation.

Biologically, relaxation oscillators appear to underlie many important cell processes. As a concrete example, let us briefly discuss the early embryonic cell cycle in frog eggs (*Xenopus* oocytes), in which there occur a set of 12 synchronous cell divisions, starting from just one cell in the fertilized egg and resulting in 4096 cells. The normal cell cycle in a mature organism involves several steps: mitosis (M, the actual cell division) and interphase, the latter made up of the substeps Gap1 (G1, when the cell grows, in preparation for cell division), synthesis (S, when DNA is replicated), and Gap2 (G2, a second gap before returning to M). In the early embryo, though, there are no checkpoints (stopping at gaps), and the cell divisions take place in quick succession. The divisions are controlled by proteins named *cyclin-dependent kinases (Cdk's)*, so called because they are active when cyclins (another type of protein) are bound to them. Examples of Cdk's are *cell-division cycle (Cdc)* proteins. A dimer made up of one of these, Cdc2, and cyclin B, a type of cyclin, is called *mitosis promoting factor (MPF)*. MPF can be in four different phosphorylation states, depending on the binding at the amino acid in position 167 (which is a threonine, and hence is referred to as "threonine-167") by the protein kinase CAK, and at a tyrosine-15 site by another protein called Wee1 when the latter is non-phosphorylated. The phosphorylation of MPF at the tyrosine-15 site is reversed by yet another protein called Cdc25, which acts in that manner when it is phosphorylated. The active form of MPF is that in which *only* the threonine-167 has been phosphorylated. When MPF is active, it phosphorylates both Wee1 and Cdc25. Leaving aside, for simplicity, the action of CAK and two of the phosphorylation states, the reactions between MPF, Wee1, and Cdc25 are as shown in Figure 29. (See [58] for more details.) This system is a net positive feedback system.

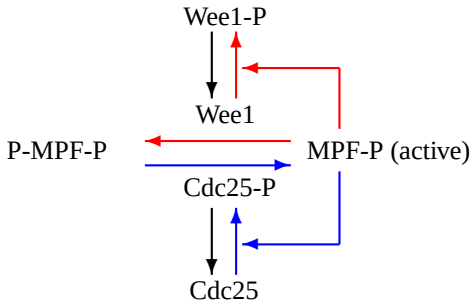

Figure 29: Cell cycle subsystem

One can give an ordinary differential equation model, and appropriate parameter values, so that there are two possible steady states: one corresponding to mitosis (high concentration of activated MPF as well as Wee1-P and Cdc25-P) and one to interphase (higher concentration of inactivated MPF, Wee1, and Cdc25). Depending on the total amount of MPF (adding active and inactive forms), these two states may theoretically exist in the same system (bistable regime) or only one of them may be possible.

It is believed that the oscillations are produced by a relaxation oscillation mechanism: the concentration of cyclin B can be viewed as a parameter which controls the concentration of MPF. Through a negative feedback loop involving yet other players (not shown), cyclin B is degraded when MPF is activated, making the system move between the monostable and bistable regimes, much as with the light-switch example. How does one test this hypothesis? In a beautiful experimental demonstration, Joe Pomerening and Jim Ferrell at Stanford blocked the degradation of cyclin B by introducing instead a mutated form which cannot be degraded. In effect, this broke the negative feedback loop and left the system in Figure 29 isolated. To verify that this system is indeed bistable, they manipulated the concentration of cyclin B and let the system relax to steady state. If the system is indeed bistable, a bifurcation diagram like the one shown in the right panel of Figure 27 should result. Indeed, the results, shown in the Western blots in Figure 16, indicate just this hysteretic behavior ("going up" versus "coming down" in parameter space). Further confirmation of this bistable behavior was obtained

from morphological observations. Figure 30 shows, for different values of the parameter and at steady-state, observations done under a microscope. (We omit details of how this was done, which is an interesting story in itself.) The pictures for parameter values between 40 and 60 show the two possible steady states in this bistable system, each of which is arrived at depending on the history of the system. As the parameter is slowly increased from 0 to 100, starting in interphase, we see nuclei that stay well-formed, indicating interphase, for a large range of parameters, while M phase (nuclear envelope broken down, chromosomes condensed) is only observed for the value 100. Conversely, going down, the M phase view persists for a large range of parameters.

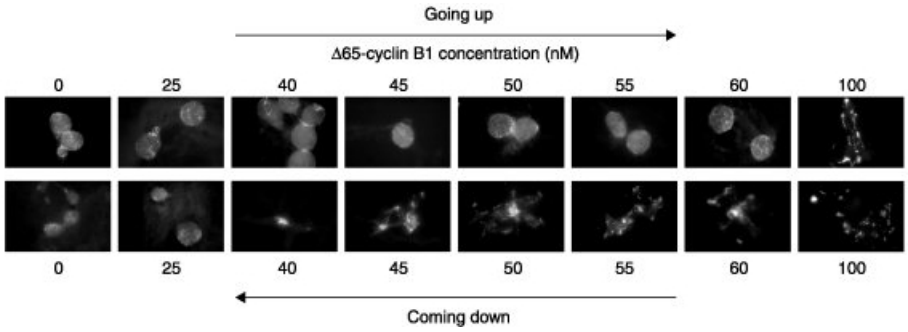

Figure 30: Hysteresis and bistability seen under a microscope, from [73]

3.2 A System-Theoretic View

In summary, many of the dynamical behaviors typical of engineering and other natural systems are of great interest when analyzing molecular biology problems, and these behaviors can be, and are, studied experimentally. This is the reason that the field has attracted the attention of experts in dynamical systems as well as in many areas of physics. On the other hand, one of the important themes in current molecular biology thought ([47, 63]) is that of understanding cell behavior in terms of cascades and feedback interconnections of elementary "modules." Cells can be seen as composed of a large number of subsystems, involved in various processes such as cell growth and maintenance, division, and death. The hope is that one should be able to decompose into such, hopefully simpler, subsystems, and then study the emergent properties of interconnections. The control and systems-theory paradigm of input/output systems, which are built out of simpler components that are interconnected according to certain rules (Figure 31) is very natural in this context, and it should permit the recursive

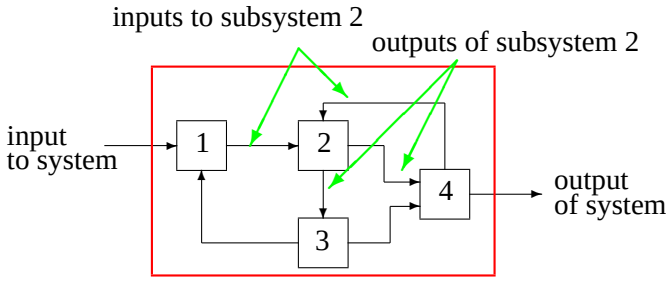

Figure 31: A System seen as an interconnection of subsystems with inputs and outputs

verification of important properties through the use of standard analysis tools such as passivity, small-gain, or input to state stability. Even if the entire system were autonomous, in order to be able to define such interconnections, one would be forced to consider subsystems that process time-dependent input signals into output

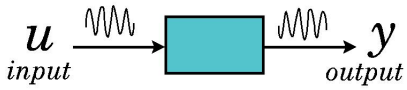

Figure 32: The systems-theory paradigm

signals (Figure 32). But, in fact, cells are not autonomous systems. They process external information, provided by physical (UV or other radiation, mechanical, temperature) or chemical (drugs, growth factors, hormones, nutrients) inputs. They also produce signals which we may view as outputs, such as chemical signals sent to other cells, commands to motors that move flagella or pseudopods, or the internal activation of transcription factors which may be monitored by measurement technologies as we have already described.

Thus, the control-theory formalism —in contrast to dynamical-systems theory, which deals with isolated systems— is not only reasonable, but natural. For example, and using ordinary differential equations for concreteness, one should study systems with inputs and outputs, in the standard sense (see e.g. [88] or any other control-theory textbook):

$$\begin{aligned} \dot{x}_1(t) &= f_1(x_1(t),\ldots,x_n(t),u_1(t),\ldots,u_m(t)) \\ &\vdots \\ \dot{x}_n(t) &= f_n(\underbrace{x_1(t),\ldots,x_n(t)}_{\text{states}},\underbrace{u_1(t),\ldots,u_m(t)}_{\text{inputs}}) \end{aligned}$$

supplemented by a selection of output variables $y_1,\ldots,y_p$ which are functions of the state:

$$y_j(t) \;=\; h_j(x(t))\,,\; j=1,\ldots,p\,.$$

The inputs, which can be seen as controls, forcing functions, or external signals, act as stimuli. Outputs can be thought of as responses, such as movement, or measurements provided by biological reporter devices like GFP that allow a partial (if $p<n$) read-out of the system state vector $(x_1,\ldots,x_n)$.

4 Challenges to Control Systems Theory

Many of the systems-theoretic questions that one would normally pose for a dynamical system such as represented by the cancer network diagram shown in Figure 9, are precisely those that leading biologists are asking, if sometimes in different language:

- What is special about the information-processing capabilities, or input/output behaviors, of such networks, and how does one characterize these behaviors?
- How do the different signal transduction pathways interact?
- How does one find the algebraic forms of reactions, and values of parameters (identification, reverse engineering)?
- Once these forms of reactions are known, how does one estimate time-varying internal states, such as the concentrations of proteins and other chemical substances, from input/output experiments (observer problem)?
- What subsystems ("modules") appear repeatedly in the same cell?

- Where do the main sensitivities affecting robustness of the system lie?
- What are the reasons (control objectives, signal processing) that there are cascades and feedback loops?
- More generally, what can one say, if anything, about stability, oscillations, and other dynamical properties of such complex systems?
- In addition to analysis questions, there are, of course, also synthesis ones, dealing with the *control* of cellular systems through drugs or genetic modifications.

Much research addresses the above types of problems for cell signaling systems, and a major and long-term research effort will continue toward their solution.

Nevertheless, I would argue that in spite of its immense success in engineering, "off the shelf" application of known control theory is not always appropriate. This is because detailed models are hard to come by: it is virtually impossible to experimentally validate the forms of the nonlinearities used in reaction terms, and even when such forms are known, to accurately estimate coefficients (parameters). New tools must to be developed in order to bridge the "data-rich/data-poor" dichotomy that exists in systems biology: relatively good knowledge of overall network structure but poor quantitative resolution. In addition, issues such as robustness, multi-scale modeling, continuous/discrete interfaces, and seamless integration of hybrid stochastic/deterministic systems, although treated to various degrees in the control field, cannot often be handled with the tools available, which were developed with very different engineering applications in mind. For the remaining part of this article, therefore, I would like to focus not so much on what existing control theory and tools can do for systems biology but rather on a sort of converse, namely how new questions in control theory arise from problems in systems biology. Even though many problems in systems biology resemble standard problems in control theory, on closer inspection they often turn out to differ in fundamental ways, and these differences are challenging and worth exploring. Let's briefly discuss this point with some examples, and later pick one particular topic for a more detailed analysis.

4.0.1 Positivity, Nonlinearities, Equilibria, Measures of Performance

An important characteristic of biochemical models is that most variables take only nonnegative values, since they represent chemical concentrations. By contrast, in the control theory literature, it is common to allow for negative as well as positive values of displacements, forces, velocities, and so on. This is especially true in linear control theory, which in essence concerns itself with differences between actual and desired values of signals. For the latter, nonnegativity makes matters technically much harder, since linear algebra techniques must be complemented by tools from convex analysis and positive linear algebra. Indeed, there is already a substantial theory of positive linear systems, motivated largely by biochemical applications. Researchers in discrete-event and several other subfields of control theory have also developed powerful techniques for dealing with nonnegativity constraints.

We already discussed how certain types of nonlinearities, such as saturations and sigmoidal responses, appear in biomolecular models. The study of systems with saturation is also routine in control theory, arising for example from actuator constraints, but the algebraic form of the saturations tends to be different: a rational function like $\frac{ap}{b+p}$ is common in biological models. Note that an expression of this form only makes sense if we know that the denominator cannot vanish. This is not a problem when $b > 0$ and p is a chemical concentration, and hence nonnegative.

Tying together the comments about positivity and about algebraic forms of nonlinearities is the meaning of seemingly obvious terms such as "negative feedback." In classical applications of control theory, a negative feedback is typically a function that takes negative values for positive arguments and vice-versa, such as the linear control law $u = K(x) = -kx$ when $k > 0$ and x and u are scalar variables. This makes no sense

for systems which are restricted to positive variables. Indeed, a negative or inhibitory feedback in biochemical applications represents more likely a function like $\frac{a}{1+kx}$. For instance, given a reaction like $\dot{p} = -p + u$ driven by an input u, an inhibitory feedback of p on itself might be represented by $\dot{p} = -p + u/(1 + kp)$: when p is small, we have the original system $\dot{p} = -p + u$, but for p large the effect of the forcing term is damped, and we have a decay $\dot{p} = -p$. As a consequence, *equilibrium locations change under feedback*. Feedback theory, for example theorems on stabilization, assume as a matter of course that the feedback is zero when we are already at a desired equilibrium, which is without loss of generality taken to be the origin of coordinates: $K(0) = 0$. However, if we look at the example $\dot{p} = -p + 1/(1 + kp)$, we see that the equilibrium location $p = p(k) = (\sqrt{1 + 4k} - 1)/(2k)$ depends on the strength of the gain k.

The word "equilibrium" brings up in itself an interesting issue. In normal mathematical usage, the words equilibrium and steady state are interchangeable. They both mean, say for a differential equation $\dot{x} = f(x)$, a point at which the rate of change is zero, i.e., a root of $f(x) = 0$. In mechanical and electrical engineering applications, typically an equilibrium is a state at which no physical or electrical activity takes place. In biological models, in contrast, where x represents the concentrations of various chemicals, a steady state is one in which the concentration is "constant" in the macroscopic sense, but this does not mean that chemical reactions are not taking place: the underlying process is stochastic and one is representing an equilibrium *probability distribution*. An equilibrium in the thermodynamic sense means something much stronger. When communicating with biologists and physicists, it is better therefore never to use the word "equilibrium" when talking about steady states.

Positivity, different nonlinearities, and the meaning of steady states, all lead one also to consider alternative norms on signals and measures of performance than are usual in control theory. See for example [49, 77, 21] for some such notions, which quantify signal duration and amplitude. Even for more classical measures, however, positivity brings up new points of view. To take the simplest possible example, consider the bilinear system $\dot{x} = -x - ux = -(1 + u)x$. The input $u(t) \equiv -2$ gives rise to $\dot{x} = x$ and hence instability. Thus, the H_∞ (i.e., induced operator L^2) norm of this system is, clearly, infinite. But suppose that we restrict ourselves to positive inputs: then $1 + u$ will remain positive, so instability is ruled out. In fact, one can explicitly compute the H_∞ norm of certain types of cascades appearing in signaling pathways, including, in particular, bilinear systems, and prove that it is finite for positive inputs; see [24, 23]. A rich, and different, theory results in this manner. What is remarkable is that, in contrast to the usual impression from hybrid and discrete event systems theory that positivity makes matters harder to analyze, in this case, one can argue that problems become easier, in some sense, due to positivity.

Another issue worth commenting on, which we expand upon below, is that of precision. In areas of bioengineering, such as anesthesia or drug dispensers, pacemakers, or prostheses, it is imperative to develop high-precision models. However, in cell modeling, models should not depend on tight values for parameters. One should avoid "pseudo-exactness," which is meaningless in that context.

The role of optimal control theory is as yet unclear. While obviously optimal control should play an important role in areas such as optimizing drug delivery, it is unclear what its role will be at a cellular-level, again because of the lack of precise measurements. One might speculate that general theoretical questions, such as whether optimal trajectories should be bang-bang or singular, are more relevant than actually solving optimization problems precisely. Another potential application is inverse optimality: if one postulates that evolution has produced at least a locally optimal design, theory might help in guessing which criteria are being optimized by a specific cellular mechanism.

4.0.2 Robustness from Structure

Few engineered systems perform acceptably under truly large uncertainties in parameters, in contrast to living cells, which perform satisfactorily in the face of large variations in intracellular concentrations of chemicals. The variability might be a consequence of unequal division among daughter cells during mitosis, which would

mean that different cells inherited different amounts of chemicals, or of genetic phenomena such as gene duplication or mutation. If functions critical to the survival of the organism are not to be affected, this means that evolution must have selected for extremely robust structures, whose study is interesting in itself and in view of its potential for suggesting novel robust designs for engineering applications.

One line of work motivated in this way is the study of stability of biochemical networks, and in particular the search for general principles that guarantee global asymptotic stability. Some of the tools developed so far in control theory are useful for the analysis of stability of biochemical networks. Others, such as those based on Hamiltonian dynamics with additional damping terms, cannot be expected to be as appropriate as they are for, say, mechanical or electrical systems. So, it is useful to develop approaches that take advantage of special properties of biochemical dynamics; two of these are as follows.

One approach to chemical network stability is based on work done in chemical engineering by Feinberg, Horn, and Jackson in the 1970s ([39, 53]). These researchers showed —under rather restrictive graph-theoretic assumptions on the structure of the chemical reactions— how to obtain local stability results as well as uniqueness of equilibria (a general global stability result was claimed in one of the papers by Horn and Jackson, but it was later retracted). The beauty of these theorems is that they are valid essentially for all parameters appearing in the system description, as long as the interconnection structure of the system satisfies certain rules and the reactions are given by polynomials (more precisely, ideal mass-action kinetics). See [89] for partial extensions to global convergence, estimates of robustness to unmodeled dynamics and results on stabilization, and an application to the stability analysis of the kinetic proofreading immunological model proposed in [68] to explain how T-cells balance sensitivity and selectivity of response in antigen recognition, as well as the observer theory in [22], and the applications to parameterizations of receptor-ligand dynamics in [25].

Another promising line of research on chemical network stability is based on the notion of monotone systems. We discuss this subject in Section 5, so we will not say more here.

Of course, stability is just one of many possible properties to be studied for robustness. More generally, one may ask how is it that a given phenotype arises, if parameters are so uncertain. An instance that has attracted much attention from the systems biology community is the work in [101]. The authors of that paper proposed that only models that display desired behaviors over relatively large ranges of parameters can be valid. They carried out a concrete study of this matter for a model of the *Drosophila* segment polarity gene regulatory network, a gene network that plays a key role in fruit fly development. Certain observable patterns in the embryo correspond in a one to one fashion with certain combinations of genes being expressed in adjoining cells, in steady state. So, given a model, one can define a set S of parameters for which the steady states that result are consistent with the biologically observed pattern. A robust model is, roughly, one for which the volume of S, as a fraction of the volume of the space of all possible parameters, is large. The authors then compared different models using this notion of robustness as a selection criterion. This way of thinking is conceptually sort of converse to robustness approaches in control theory such as the calculation of stability radii under structured or structured perturbations.

The authors of [1] argued that this large robustness to parameter variations should represent a characteristic of the network itself. As a simple illustration of the idea, suppose that, in some model, the concentration of some substance p depends on the concentrations of x and y as $p = \frac{xy}{b+xy}$, but that the actual value of the constant $b > 0$ is unimportant in the sense that the same macroscopic characteristics are observed for a large range of different values of b. Then, one may reasonably argue that all that matters is that both x and y being large results in $p \approx 1$, but that if one of them is small, then $p \approx 0$. Continuing with this line or reasoning, it is then proposed to employ a purely Boolean model, where x and y take only two values, "0" and "1" for "high" and "low" respectively, and similarly for p. The dependence of $p = p(x, y)$ is now simply the logical "and" of the two input variables. In [1], such a Boolean model is analyzed for the *Drosophila* segment polarity gene regulatory network, and biological and experimental consequences are described. The follow-up paper [20] argues that asynchrony must be taken into account, and a detailed analysis of the now-stochastic dynamics (in an asynchronous model, state updates are not deterministic) is carried out and several biological predictions are

given.

4.0.3 Alternative Modeling Frameworks

The issue of comparing continuous and Boolean models for fly development brings up another key difference from more classical control theory: the selection of an overall modeling framework for a given physical process is less obvious. In control theory, quite often an appropriate formalism is dictated by the problem being considered: discrete time when dealing with sampled data systems, continuous time for analog systems, Boolean or symbolic for digital and computer systems. The right level of modeling for biological systems is far less clear.

Biochemical systems are traditionally modeled using ordinary differential equation models, usually based on mass action kinetics or singularly perturbed versions of mass action (Michaelis Menten). If spatial localization is important, and a compartmental model is not appropriate, then diffusion effects might be explicitly considered, and reaction-diffusion PDE's should be used instead of ODE's. Whether one should use PDE or ODE models is usually clear from the problem, depending on diffusion coefficients, size of the cell, and time scale of interest. However, differential equation models for concentrations represent, by definition, averages, and therefore are only meaningful for situations in which one has a large number of molecules. On the other hand, it may well happen that there are just a few dozen molecules of a given type involved in a particular cellular process at a given time, what biologists would refer to as a *small copy number*. As current technologies make it possible to observe the activities of proteins and other cellular components in single cells, it is important to have models that accurately describe that situation and permit comparisons with experiments. Thus, a stochastic model for the actual numbers of particles may be more appropriate, see e.g. [16, 59, 75, 95].

Probabilistic evolution models in control theory tend to focus on stochastic differential equations, which are based upon the paradigm of a deterministic system driven by *external noise*. For the analysis of small copy number biochemical systems, the appropriate formalism is that of Chapman-Kolmogorov equations in continuous time, or *master equations*, which track the births and deaths of units of each participating chemical species. One refers to this stochasticity as *internal noise*. Stochastic ODE's, or as physicists call them, Langevin equations, sometimes can be used as approximations for statistics (means, variances) of solutions of master equations, but the basic model is the latter. A variation of this theme is that, when particles of several different types interact, and individuals must be tracked in their behavior, agent-based models may be called for.

4.0.4 Prediction of Internal Structure

An important set of results in mathematical systems theory deals with the prediction, based on behavior, that certain structures must necessarily be embedded in systems. Having this insight could be helpful to the biologist, since it may suggest a search for particular subsystems, and, conversely, the absence of which might provide evidence that a biological entity being studied could not possibly behave in a certain fashion.

Let us take as an illustration *E. coli* chemotaxis, which we introduced earlier. As we explained, chemotaxis allows the bacterium to search for food, by evaluating a nutrient concentration gradient. In order to do so, it is important for the bacterium to be able to distinguish nonzero gradients (which indicate progress toward or away from a food source) versus constant concentrations. Thus, it must be able to somehow ignore constant gradients, while reacting to changes. Indeed, it is known experimentally from measuring responses to changes in nutrients that this response exhibits approximately a zero DC gain. The left panel of Figure 33, from [78], shows the impulse response of CCW activity of flagellar motors averaged over a population and many experiments, together with a sum-of-exponentials fit, and strongly suggests a zero DC gain; indeed; the right panel shows the experimentally measured step response, also a population average, together with (thinner curve) the integral of the impulse response. (The authors also studied how mutations in CheR, CheB, etc, would affect this behavior.) One is thus led to look for some sort of integral control mechanism in order to account for this "step disturbance

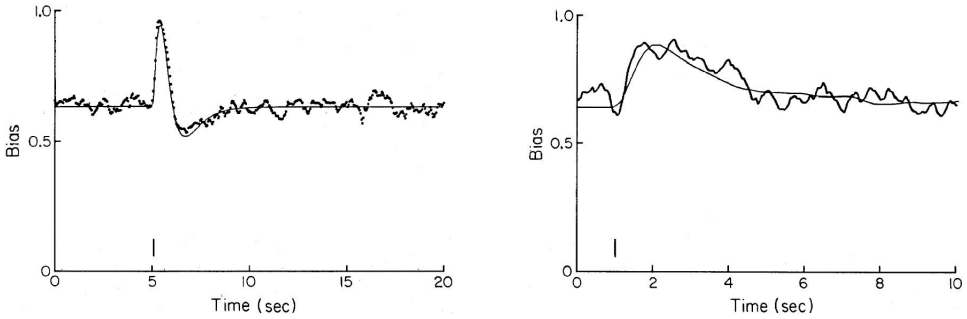

Figure 33: *E. coli* chemotactic impulse and step response, from [78]

rejection" property. Motivated by [4], the papers [104] and [55] have pursued this question, focusing on what biochemical mechanisms might implement the integrator.

An interesting set of new theoretical questions arises in this context. Recall that the classical Francis-Wonham Internal Model Principle (IMP) tells us that if a controller regulates a system against external disturbances in some family —such as steps,— and if this regulation is structurally stable in a precise mathematical sense, then the controller must necessarily contain a subsystem —such as an integrator— that can itself generate all such disturbances, and which is driven by a suitable error signal. But there are some potential drawbacks when attempting to use this theorem in biological applications. For one thing, it is not at all obvious that one may distinguish between a "controller" and a "plant" in a cell, as needed for the theory, nor is it obvious why structural stability in the precise mathematical sense required by the Francis-Wonham Theorem should hold. In addition, there is another characteristic that distinguishes the more classical engineering disturbance rejection setup from the problem of interest in cell biology: *signal detection* is of great importance. The cell must detect the signal before regulating —in fact, a *large overshoot is often desirable*, in order to "start up" a downstream subsystem. (On the other hand, adaptation after the signal is processed is also often important: the reaction might be metabolically too expensive to be kept "on" for very long). One engineering analogy would be as follows. Suppose that we wish to design an automobile active suspension system with the following additional property: if a large gradient is felt by the car –such as might happen when the driver falls asleep and starts to run off the road– some signal should be provided to the driver in order to alert him/her to the presence of a problem. This would be in contrast to minor surface roughness of roads, which should indeed be filtered out, as in more standard active suspension systems. One is thus lead to a new problem of "disturbance rejection with signal detection" which motivated the paper [91], where one can find an internal model theorem that does not require the assumption of structural stability, nor an a priori requirement for the system to be partitioned into separate plant and controller components. In lieu of structurally stable regulation, a signal detection criterion, expressed through zero-dynamics and relative degree for nonlinear systems, was imposed in order to force the existence of an internal model. Much work remains to be done on this subject, particularly in understanding approximate disturbance rejection and its trade-offs with signal detection; the paper [91] barely scratched the surface of the question.

4.0.5 Realization and Identification Questions

A very active area of research in systems biology is that of reverse engineering of gene and protein networks. Based upon measurements of concentrations of cellular species such as proteins or RNA, and their changes in response to probing inputs, the goal is to unravel the internal structure of the system (Figure 34). This topic seems like a perfect target for systems identification and realization theory techniques, and there is indeed substantial work being carried out in that direction.

However, in contrast to standard formulations of control-theoretic systems identification, in biological prob-

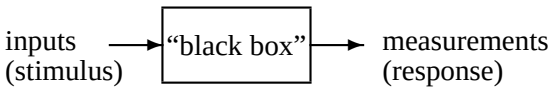

Figure 34: The reverse engineering (identification/realization) problem

lems it is often extremely expensive, if not impossible, to apply complicated "test signals." Often, biologists are only able to consider step inputs, or perhaps a small number of combinations of step inputs and ramps. Although sufficient for local identification of linearized models, this certainly does not yield enough information to characterize the behavior of a nonlinear system. The standard theory, see e.g. [88] for linear systems and [87] for a class of nonlinear dynamics, requires one to allow the consideration of arbitrary concatenations of inputs, which is not possible. Thus, one interesting question is to quantify the amount of experimentation needed in order to identify a general nonlinear system if only a finitely-parametrized family of inputs is available. The paper [90] looked into this problem and established that the best possible answer, assuming exact measurements, is $2r + 1$ experiments, where r is the number of parameters. Moreover, in a precise mathematical sense, a generic set of such experiments suffices. This resembles, but differs technically in a fundamental way, from similar results for embeddings of chaotic systems or for generic observability. An exploration of these techniques in the modeling of catabolic repression for glutamine/glutamate substrate selection in *S. cerevisiae* can be found in [99].

A related set of questions is motivated by the fact that often only *steady state* measurements are available. Let $x(t)$ be a state vector that lists the time-dependence of the concentrations of certain proteins, mRNA, or other substances, at time t, and let p be a parameter vector that represents the concentration levels of certain enzymes that are kept at a constant value during a particular experiment. Suppose that an experiment can measure the concentrations $x(t)$ after a long enough period, presumably long enough so that the system has relaxed to a steady state. Several experiments are carried out, for various different particular values of p. Let's assume that the dynamics of the system are described by a set of ordinary differential equations $\dot{x} = f(x, p)$, where the vector of state variables evolves in some manifold M, and also that for each p in a neighborhood of some default value p_0 there is a unique steady state $\xi(p)$, that is, a solution of $f(\xi(p), p) = 0$, and that the map $p \mapsto \xi(p)$ is smooth and describes a manifold. It is often the case that the equations defining the system —that is, the functions f_i describing the vector field— are unknown, even in general form, but one wishes nonetheless to determine the interaction graph of the system, that is to say, to know which variables directly influence which variables, as well as the relative strengths of these interactions. Using perturbations of parameters, the map ξ induces perturbations that are tangent to the manifold of equilibria. Formally, the possible vector fields Y in parameter space induce, by push-forward, vector fields $X = \xi_*(Y)$ on the state space, which in turn determine a distribution $\mathcal{D} := \operatorname{span}\{X\} \subseteq TM$. Under suitable genericity conditions, $\dim \mathcal{D} \equiv n - 1$, and with the main technical assumption described in [61], one can obtain that $df_i \in \mathcal{D}^\perp$ ($\langle df_i, \mathcal{D}\rangle = 0$) for appropriate coordinates of the motion, which determines each df_i projectively. That is, the interaction graph is completely characterized, and strengths are identified up to a constant common multiple. See [61] for details, as well as [93] for further results, some elementary considerations about noise in [5], and a computational complexity analysis of approximate methods for large-scale problems in [17].

5 A Qualitative/Quantitative Approach

We now turn to a specific topic in more detail. It is common in systems biology to set up a model based on biological knowledge, estimate parameter ranges, and then explore the spaces of parameters and initial conditions by means of a huge number of simulations. This is often combined with bifurcation analysis and, if applicable, model reduction techniques. There are several shortcomings to such an approach, however. First,

the *form* of nonlinearities often cannot be well-justified. Also, parameters such as reaction rates are based on rough guesses or on data from different and perhaps inconsistent sources, and usually are obtained from *in vitro* experiments (that is, experiments in a test tube or petri dish, or more generally not from an entire organism, as opposed to *in vivo* measurements). In addition, parameter and state spaces are of high dimension, which makes convergence of numerical techniques questionable and at best very local. Some of these problems are intrinsic, and cannot be solved by better technology or algorithms; for example, parameters such as enzyme concentrations vary from cell to cell, even among cells of the same type. In addition, a purely numerical approach does not provide fundamental understanding. This argues for the desirability of approaches which, while taking advantage of the huge, and growing, amount of qualitative network "schematic" knowledge such as shown in Figures 8 or 9, take into account the uncertainties inherent in biological measurements, and effectively integrate relatively sparse quantitative data. We describe in this section one such approach, based upon the systems theory paradigm of I/O systems and combining information on network structure with steady-state step response data on subsystems. A key concept will be that of monotone systems. Monotone subsystems as components of a larger system behave, for many purposes, just as if they were one-dimensional systems (single species), even though they may in fact have an arbitrarily large dimension (number of species).

Our main themes for the remainder of this paper may be summarized as follows:

• Network structure (qualitative knowledge) is important as a constraint: for example, certain structures are not compatible with oscillatory behavior.

• On the other hand, parameters matter, as they underlie bifurcation phenomena: as examples, Hill coefficients and other constants determine a hyperbolic vs. sigmoidal shape, and the gain g of a feedback loop, as discussed earlier, determines mono-stable vs. multi-stable behavior. This is a problem because such parameters are often difficult to obtain.

• The interplay of structure and parameters can be fruitfully studied by breaking up systems into well-behaved building-blocks —monotone components, for example— and using only a *restricted amount* of input/output quantitative data —step responses, for example– for these subsystems in order to characterize global behavior.

5.1 Consistent Graphs and Monotone Systems

Suppose given a directed graph G whose edges are labeled "$+$" or "$-$" (in diagrams, we sometimes use activating arrows $\rightarrow$ for positive edges and the inhibition symbols $\dashv$ for negative edges). By a path we will mean any nontrivial sequence of nodes connected by an edge. We ignore direction in this definition: for example, 1-2-3 and 3-2-1 are both paths in the graphs shown in Figure 35. A cycle is, in particular, a path with same

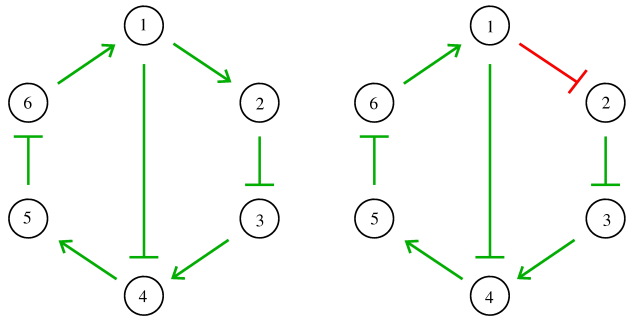

Figure 35: A consistent and an inconsistent graph

first and last node, such as 1-2-3-4-1 in Figure 35. The parity of a path is the product of the signs along that path; for example, in Figure 35, the path 1-4 has negative parity in both panels, whereas the path 1-2-3-4 has negative parity in the graph shown in the left panel but has positive parity, since there are two negative links, in

the right panel. We call a graph *consistent* or *coherent* if every possible path between any two given nodes has same parity. The graph in the left panel of Figure 35 is consistent. Equivalently, any cycle must have positive parity. A graph is *in*consistent if some pair of nodes has two connecting paths with different parity, as in the right panel of Figure 35 (compare 1-4 with 1-2-3-4).

Observe that if a graph is inconsistent, one may always pull out enough edges so that the remaining graph is consistent, and thus one may think of the original system as a "negative feedback loop" around a consistent one, the deleted connections constituting the feedback, Figure 36. For the graph in the right panel of Figure 35,

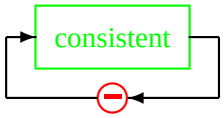

Figure 36: Pulling out inconsistent connections

for example, we could simply remove the edge 1-2.

Now consider a system of ordinary differential equations $\dot{x} = f(x)$, with no inputs nor outputs for the time being. We assume that the system is sign-definite: for each two components x_i and x_j, either x_i always inhibits x_j or x_i always activates x_j, as in Figure 19. As we argued earlier, this is not an unreasonable restriction. Mathematically, sign-definiteness means that $\frac{\partial f_i}{\partial x_j}(x)$ does not change sign as a function of x, for each pair of distinct indices i and j, where f_i denotes the ith component of f. As when defining sign-consistent graphs, we do not care about the diagonal terms $\frac{\partial f_i}{\partial x_i}(x)$ of the Jacobian of f. To any sign-definite system in $\mathbb{R}^n$ we may associate an incidence graph G on the nodes $\{1, \ldots, n\}$, where we draw an edge from node j to node i if $\frac{\partial f_i}{\partial x_j}(x) \not\equiv 0$, and assign a positive sign to this edge if $\frac{\partial f_i}{\partial x_j}(x) > 0$ for some x, and negative otherwise.

5.1.1 Monotone Systems: A More General Notion

Systems whose incidence graphs are consistent are examples of *monotone* systems. A *monotone* system is one for which there is some partial order in the state space so that the evolution operator preserves the order. Denoting the order by "$\leq$" this means that

$$x(0) \leq y(0) \;\Rightarrow\; \varphi(t, x(0)) \leq \varphi(t, y(0))$$

for all $t \geq 0$, where we are denoting by $\varphi(t, \xi)$ the solution at time t of the initial value problem $\dot{x} = f(x)$, $x(0) = \xi$, and we assume for simplicity that solutions are unique and defined for all $t \geq 0$. (Note that this definition makes sense in both continuous and discrete time.) An example of a partial order in $\mathbb{R}^2$ is the "Northeast" order, in which we declare that $(x, y) \leq (x', y')$ provided that both $x \leq x'$ and $y \leq y'$, and more generally for every n, $x \leq y$ provided that $x_i \leq y_i$ for each $i = 1, \ldots, n$. More generally, one can define partial orders associated to any possible orthant in $\mathbb{R}^n$, for example in $\mathbb{R}^2$ the "Northwest" order: $(x, y) \leq (x', y')$ provided that both $x \geq x'$ and $y \leq y'$, i.e., $(x', y') - (x, y) = (x - x', y - y') \in K$, where K is the second quadrant $\{(a, b) \,|\, a < 0, b > 0\}$. A system with a consistent incidence graph is monotone with respect to some such order: in each connected component of the graph, just pick one node N, label it $+$, and assign to any other node M in the same connected component the sign of a path from N to M. In this way, an assignment of signs to nodes is obtained, and the system can be easily shown to be monotone with respect to the order associated to the corresponding orthant. (In abstract terms, "integrals are independent of the path" implies that there is a well-defined "potential.") Another way to say this is as follows: monotonicity with respect to an orthant order is equivalent to the property that, up to possible sign-flips $x_i \leftrightarrow -x_i$ for some of the variables x_i, the system is "cooperative," i.e. monotone with respect to the "Northeast" order. Interestingly, the same procedure, now interpreting sign-flip as "$0 \leftrightarrow 1$," gives the property that is also called "monotonicity" for Boolean networks (for which see [84] and references there).

Under an additional hypothesis of *irreducibility* (basically, strong connectedness of the incidence graph plus a mild condition, such as asking that non-identically zero Jacobian entries be almost everywhere nonzero), one obtains what are called *strongly* monotone systems: $x(0) \leq y(0)$ but $x(0) \neq y(0)$ implies that $x(t) = \varphi(t, x(0)) < \varphi(t, y(0)) = y(t)$ for $t > 0$ in a strong sense which we will not define here in general, but which for systems monotone with respect to orthants amounts to $x_i(t) < y_i(t)$ for every coordinate $i = 1, \ldots, n$. What matters is that strongly monotone systems are very well-behaved in a dynamical sense. According to a beautiful result of Moe Hirsch (cf. [51, 50, 83, 52]), almost every bounded solution of such a system converges to the set of equilibria. By "almost any" one means every solution except for a measure-zero set of initial conditions, or, in a different version of the theorem, every solution except for those starting from a thin set in the Baire category sense. In particular, no chaotic or other "strange" dynamics can occur; in fact, not even limit cycles can arise in strongly monotone systems.

The intuition behind this result is particularly clear for two-dimensional systems

$$\begin{aligned} \dot{x} &= f(x, y) \\ \dot{y} &= g(x, y) \end{aligned}$$

which are mutually activating, or as is said in mathematical biology, *cooperative*: $\frac{\partial f}{\partial y} \geq 0$, $\frac{\partial g}{\partial x} \geq 0$. Let us compare any two states (x, y) and (x', y') such that $x' = x$ and $y' > y$, and assume that $f(x, y) \geq 0$ (i.e., the trajectory is moving right) at that state; see Figure 37. Now, since $\frac{\partial f}{\partial y} \geq 0$, it follows that $f(x', y') = f(x, y') \geq$

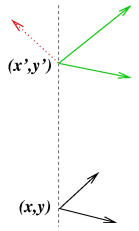

Figure 37: Only possible motions at (x', y') are rightward

$f(x, y) \geq 0$, *so* (x', y') *must also move rightward*. In other words, the dotted arrow is impossible. This implies that there cannot be a periodic orbit. Otherwise, and assuming for definiteness that there is a periodic orbit that is oriented counterclockwise, there would exist on such an orbit two points as shown in Figure 38, contradicting

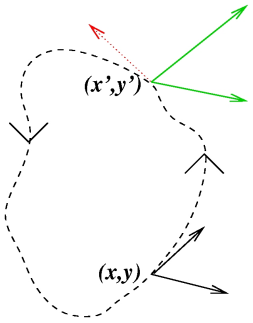

Figure 38: No possible periodic orbits

the above argument because the motion at (x', y') is tangent to the curve and thus would have to look like the impossible dotted arrow. (The clockwise case can be argued similarly, using that $\frac{\partial g}{\partial x} \geq 0$.)

5.1.2 Monotonicity Might not be Obvious

Often in applications, a system that is not monotone as originally modeled turns out to be so under some simplifications. An elementary illustration of this phenomenon is as follows. Suppose that an enzyme B catalyzes conversion of C to A, as in the left panel of Figure 39. (We also included a reverse conversion reaction, which is independent of B, since such reactions are usually present; however, a reverse reaction is not required for the point to be made.) Thus, B negatively affects the concentration of C, and positively

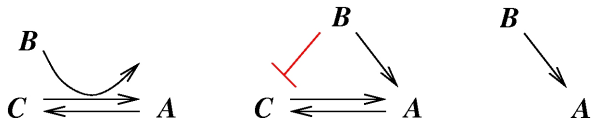

Figure 39: Consistency after elimination

affects that of A. It we introduce lower case variables $a(t), b(t), c(t)$ for the concentrations of the three species, under various simplifying assumptions a set of ideal mass-action equations for this reaction would be $da/dt = k_1bc - k_2a$, $db/dt = 0$, and $dc/dt = -k_1bc + k_2a$. (More complicated equations would involve Michaelis-Menten dynamics, but this simple set of equations suffices to make our point.) The incidence graph for this set of differential equations will look as the middle panel of Figure 39, and thus be inconsistent. One would then think that monotone theory cannot be applied to this example. However, it is clear from the equations –and natural, since they are merely two forms of the same protein– that the sum of the concentrations of A and C must be constant in time, i.e. $d(a+c)/dt \equiv 0$, so there is a conservation law $a(t)+c(t) \equiv c_0$ constant (we may think of c_0 as the initial amount of C, if $a(0) = 0$). Geometrically, the hyperplanes $a + c = c_0$ (intersected with the non-negativity condition $a \geq 0$ and $b \geq 0$) are invariant manifolds. Thus, one may eliminate c (or a) from the system of differential equations, leading to $da/dt = k_1b(c_0 - a) - k_2a$, $db/dt = 0$, and this reduced system is now consistent, since there are no loops (remember self-loops are ignored), cf. the right panel in Figure 39. To analyze solutions of the differential equation, we may first restrict to an appropriate hyperplane, which depends on the initial conditions, and monotone theory can therefore be applied. This observation is key in applications to signaling cascades, where A and C might correspond to un-phosphorylated and phosphorylated forms of the same protein, for example. The MAPK cascade example treated later is a less trivial illustration of this same idea. *Very often, much less obvious eliminations and changes of coordinates are needed in order to be able to use monotone theory to analyze a given reaction, and the search for such transformations is an active area of research*, see for example [100, 30, 25, 32, 12, 11, 10, 31, 8, 7].

Moreover, as always in applied mathematics, one must pick a level of modeling appropriate for mathematical tractability. Just as, in elementary mechanics, one assumes that springs are linear, it may be useful to drop weak connections –or perhaps, reactions occurring at a faster time-scale– and first analyze a system as if it were monotone. In any event, the extension to *input/output monotone* systems, discussed next, allows one to build large monotone systems out of smaller components, and also, more interestingly, to fruitfully study non-monotone systems, at least in some cases, by viewing them as negative feedback loops around monotone systems.

5.2 I/O Monotone Systems

A continuous-time finite-dimensional system, in the standard sense of control theory:

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

(one may also study delay-differential systems, reaction-diffusion PDE's, and more abstract flows in metric spaces, cf. [36]) is *monotone* if there are nontrivial orders in the state, input, and output spaces, such that

$$\xi_1 \leq \xi_2 \quad \& \quad u_1 \leq u_2 \quad \Rightarrow \quad x(t, \xi_1, u_1) \leq x(t, \xi_2, u_2) \quad \forall\, t \geq 0$$

with respect to the state and input orders, and the output map h preserves the order as well. As usual, $x(t, \xi, u)$ means the solution at time t if the initial state is ξ at $t = 0$ and the input is $u(\cdot)$; and $u_1 \leq u_2$ for controls means that $u_1(t) \leq u_2(t)$ for all t. When there are no inputs nor outputs, this reduces to the earlier definition of monotone systems. The generalization to I/O systems is from [12, 11], and it was motivated by the types of problems that we are discussing here. Orders are typically defined by positivity cones K, by defining $\xi_1 \leq \xi_2$ to mean $\xi_2 - \xi_1 \in K$, and similarly for input and for output values. For cones, monotonicity can be checked in infinitesimal terms, not requiring solution of differential equations. A very special but most important case is that of monotonicity with respect to cones that happen to be orthants in Euclidean space. Suppose that a system is sign-definite, meaning that we can draw unambiguous sign-graphs for the Jacobians of f and h, analogously to what we did for systems with no inputs nor outputs. More precisely, $(\partial f_i/\partial x_j)(x, u)$ has a constant sign $\varepsilon_{ij} \in \{0, +, -\}$ for all (x, u) and all $i \neq j$ (we may ignore self-loops), and, for all i, j and (x, u), $(\partial f_i/\partial u_j)(x, u)$ has a constant sign $\alpha_{ij} \in \{0, +, -\}$ and $(\partial h_i/\partial x_j)(x)$ has a constant sign $\beta_{ij} \in \{0, +, -\}$. A system is monotone with respect to *some* orthant if and only if its incidence graph does not contain any negative cycles.

For the monotone system

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

consider step, i.e. constant, inputs $u(t) \equiv u$. One can prove, under weak boundedness assumptions, that for for each u, there is at least one steady state: $f(x, u) = 0$, and that for each periodic input $u(t + T) = u(t)$, there is a corresponding periodic solution (which is a limit cycle, if unique). Now assume that for each such constant input it holds that all solutions are bounded (this is frequently the case in biological systems, due to conservation laws), and that there is a *unique* steady state x_u corresponding to this value of the input,. Under weak additional hypotheses ([56, 27]), one can the prove that x_u must be a global attractor, i.e. all solutions of $\dot{x} = f(x, u)$ converge to x_u as $t \to \infty$. We say in this case that the system has a *monostable steady-state step response* and define (composing with the output map) the *characteristic* or *steady state step response* of the system as the map $u \mapsto k(u) := h(x_u)$; see Figure 40.

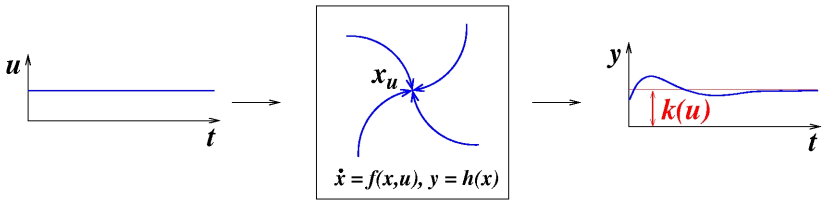

Figure 40: Characteristic

Monotone systems with well-defined characteristics constitute a very well-behaved set of building blocks for arbitrary systems. In particular, cascades (Figure 41) of such systems inherit the same properties. Moreover,

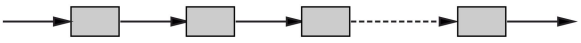

Figure 41: Cascades

there are asymptotic gain estimates: the omega-limit sets satisfy $k(\liminf u(t)) \leq \Omega^+[x(t, u)] \leq k(\limsup u(t))$ for any $u(\cdot)$, and in particular if $u(t) \to \bar{u}$, then the output satisfies $y(t) \to k(\bar{u})$, see Figure 42.

5.2.1 Blending Qualitative and Quantitative Data

The most important fact in the present discussion is that characteristics *can often be measured experimentally*. We discussed this fact in Section 3.1, and illustrated in Figure 24. This is in contrast to actual system parameters,

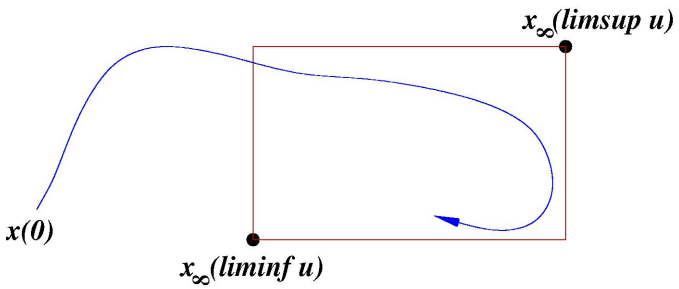

Figure 42: Asymptotic gains

which are typically hard to estimate. *In our approach, we blend qualitative information about the system, specifically monotonicity, with quantitative information, specifically characteristics*. We view this as one way to bridge the "data-rich data-poor" gap, but we are also confident that other approaches, most probably totally unrelated to monotonicity and characteristics, will be developed in the future to similarly combine qualitative and quantitative data.

5.3 Main Theorems

Some of the main results for monotone I/O systems characterize the location and stability of equilibria of closed-loop systems. The results represent nontrivial generalizations of the elementary facts that we discussed in connection with Figure 27.

For simplicity (but see generalizations in [37]), we assume from now on that inputs and outputs are scalar, i.e. $m = p = 1$, and that the order in the input and output value spaces is the usual one in $\mathbb{R}$. The main positive-feedback theorem is from [11], and is as follows. Suppose given a system $\dot{x} = f(x,u)\,,\ y = h(x)$ which is monotone and admits a characteristic k. We wish to study the closed-loop system obtained when using a feedback law $g : y \mapsto u$ that is monotone increasing, as depicted in Figure 43. (More generally, an

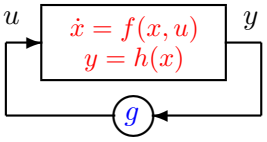

Figure 43: Positive feedback configuration

entirely analogous result holds if g is replaced by the characteristic of a monotone dynamical system.) We now plot k together with the inverse of g, and label each intersection between the two graphs by an "S" or an "U" depending on whether the slope of the graph of k is smaller ($k'(u) < (g^{-1})'(u)$), or larger respectively, than that of g^{-1}. See Figure 44 (We assume that the graphs intersect transversally.) The conclusion, under a nondegeneracy assumption of strong monotonicity of the closed-loop, is that steady-states of the closed-loop system are in a one-to-one correspondence with the intersections, and that almost every bounded trajectory (with the possible exception of a set of measure zero of initial conditions that give trajectories approaching saddle points associated to U's) converges to a steady state associated to one of the S's.

A monotone system under monotone feedback is still monotone, so this result is one about monotone systems. The main point is that *conclusions about the closed loop system, which may have arbitrarily large dimension, are derived from looking at a simple one-dimensional picture.*

It is particularly interesting that *the result remains true even if arbitrary delays are allowed in the feedback loop* (infinite phase margin) *as well as if diffusion is added*. By the latter statement we mean that for the

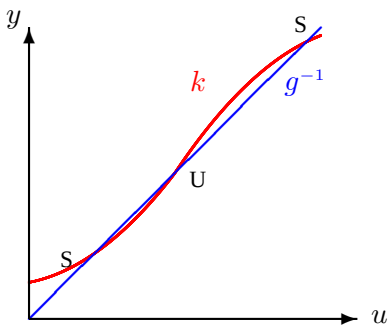

Figure 44: Intersection of characteristic and positive feedback

cooperative reaction-diffusion equation

$$\frac{\partial x}{\partial t} = D\Delta x + f(x,u)$$

for $x = x(t,q)$ with q belonging to a convex domain and with no-flux (Neumann) boundary conditions, diagonal diffusion matrix D, and assuming a discrete set of equilibria, almost all solutions converge to one of the uniform states predicted if the corresponding ODE is analyzed. In other words, *no Turing-like pattern formation due to diffusive instability can occur.* This follows by combining the above theorem with deep results due to Kishimoto, Weinberger, Casten, Holland, and Matano, as discussed in [92].

As we discussed earlier, non-monotone systems can be viewed as "negative feedback loops" obtained by pulling-out inconsistent interconnections from the original system. Let us discuss next a negative-feedback result, again, for simplicity, restricting ourselves to the case $m = p = 1$ (see [36] for generalizations to multiple inputs and outputs). The result, from [12], is as follows. Suppose given a system $\dot{x} = f(x,u)\,,\ y = h(x)$ which is monotone and admits a characteristic k. We again wish to study a closed-loop system as in Figure 43, but now assuming that the feedback law $g : y \mapsto u$ is monotone decreasing. We once more plot k and g^{-1}. The intersection between the plots, if it exists, is necessarily unique, see Figure 45. We consider the following scalar discrete time iteration:

$$u_{i+1} = (g \circ k)(u_i)\,.$$

The result is that, if this iteration has a globally attractive fixed point $\bar{u}$ (as shown in Figure 45), then the closed-loop system, provided that trajectories are bounded, has a globally attracting steady state.

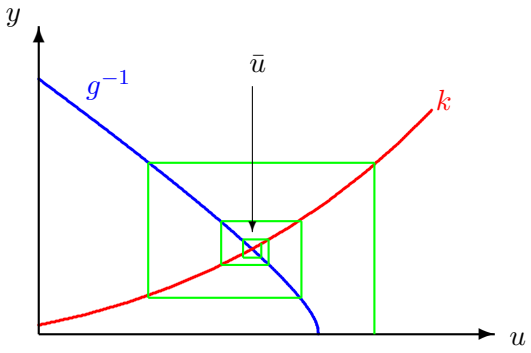

Figure 45: Intersection of characteristic and negative feedback

As with the positive-feedback case, this *result remains true if arbitrary delays are inserted in the feedback loop, and if diffusion is allowed.* (The diffusion theorem was only recently proved, in [35]. It is very surprising, because the closed-loop system is not monotone, and thus standard comparison theorems from PDE theory

cannot be used directly.) We remark that for a linear monotone system, the H_∞ gain is maximized at zero frequency, and the result would follow from the Nyquist criterion. Thus, we think of this result as a small-gain theorem. For applications of the negative feedback monotone system result to species competition problems see [32], to circadian rhythm models in *Drosophila* see [9], and to a model of testosterone dynamics see [38].

The positive and negative feedback theorems just discussed are closely related to the general notion that negative loops are required for oscillations, and nontrivial positive feedback for possible multiple steady states, [96, 86, 26, 45]. They also constitute a generalization of some aspects of the beautiful and deep work carried out by Tyson, Othmer, Mallet-Paret, Smith, and others, starting in the mid 1970s [76, 48, 3, 72, 98, 97, 65, 42, 83, 82], for inhibitory or activating cyclic structures $\dot{x}_1 = f_1(x_n, x_1)$, $\dot{x}_2 = f_2(x_1, x_2)$, $\ldots$ $\dot{x}_n = f_n(x_{n-1}, x_n)$ (scalar x_i's), as in Figure 46. These systems were motivated by the "Goodwin model" of gene expression

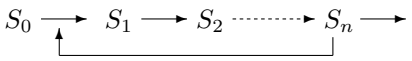

Figure 46: Cyclic systems

and similar systems in the mathematical biology literature (e.g. [70, 44]); see also the recent related passivity-based results in [13]. The papers [65, 42, 83, 82], among other major results, established a Poincaré-Bendixson theorem which tightly characterizes Ω-limit sets for such systems, in terms of periodic orbits and heteroclinic connections among equilibria.

5.4 Almost-Monotonicity

The approach that we have been discussing is based on the decomposition into interconnections of monotone subsystems. When using tools like the small-gain theorem, the smaller the number of subsystems, the easier to apply the theory. We already mentioned that systems might be monotone after appropriate transformations. On the other hand, obviously other systems will not be, even after simplifications. Let us call the smallest number of edges that must be removed in order to obtain a consistent graph the *consistency deficit (CD)* of the graph. As an example, consider the graph in Figure 47. The CD is 1, since it is enough to just remove the diagonal positive

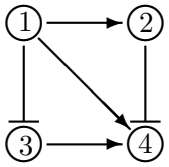

Figure 47: Graph with CD $= 1$

edge. (If one had only allowed removal of inhibitory edges, then both 1-3 and 2-4 would have to be taken out to achieve consistency.) Computing the CD exactly for large-scale graphs is not trivial: Using a reduction to max-cut, [28] shows that the problem is NP-hard. However, a polynomial time approximately-optimal algorithm is given there for computing CD, using semi-definite programming techniques.

Informally, let us say that a graph is *almost-consistent* —or an associated dynamical system *almost monotone*— if the CD is small compared to the original number of edges in the graph. Using this terminology, it has been pointed out ([69, 66, 67]) that almost-consistent biological circuits seem common. The work [64] examines the *E. coli* transcriptional regulation map provided by Uri Alon's lab (577 interactions between 116 transcription factors and 419 operons) and estimates a much smaller CD than for a randomized version of the same network. These preliminary results provide a strong indication that almost-consistency is ubiquitous in biological networks. In the same large-network statistical analysis spirit, one may ask if smaller CD is correlated

with more ordered (less "chaotic") behavior. It is hard to perform this type of analysis on differential equations, but for Boolean networks, the paper [85] described efficient ways to compute Derrida curves ([57]), which are a discrete analog of Lyapunov exponents. If the answer is positive, this would provide motivation for the general approach of studying systems that are close to being monotone. Work is in progress in this direction.

To conclude this paper, let us discuss one example, which is, in fact, the example that started this line of research and led to the definition of I/O monotone systems.

5.5 MAPK Cascades

Mitogen-Activated Protein Kinase (MAPK) cascades are a ubiquitous "signaling module" in eukaryotes, involved in proliferation, differentiation, development, movement, apoptosis, and other processes ([54, 62, 102]) Figures 8 and 9 show MAPK cascades as subsystems. There are several such cascades, as diagrammed in Figure 48, and they share the property of being composed of a cascade of three kinases, see Figure 49 (see [19] for

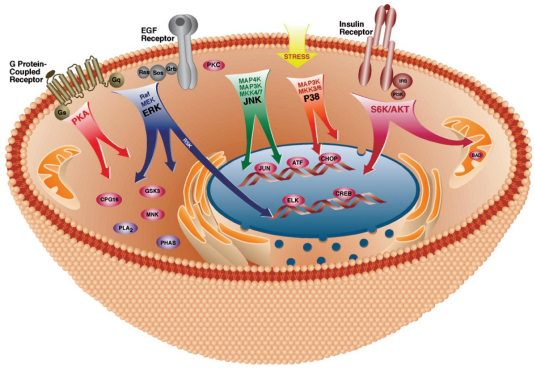

Figure 48: MAPK Cascades, reprinted by permission from [79]

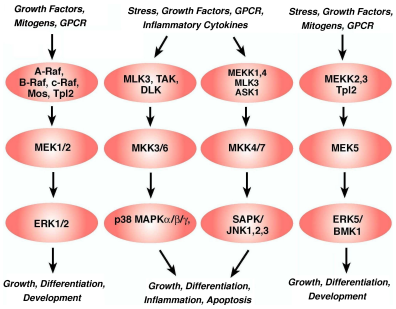

Figure 49: Different MAPK cascades, similar structure.

several similar illustrations). The basic rule is that two proteins, called generically MAPK and MAPKK (the last K is for "kinase of MAPK," which is itself a kinase), are active when doubly phosphorylated, and MAPKK phosphorylates MAPK when active. Similarly, a kinase of MAPKK, MAPKKK, is active when phosphorylated. A phosphatase, which acts constitutively (that is, by default it is always active) reverses the phosphorylation. There are many models of MAPK cascades, with varying levels of complexity. We base our discussion upon the Huang-Ferrell model ([54, 8]), see Figure 50, and use $z_1(t)$ for MAPK concentration, $z_2(t)$ for the concentration of the singly-phosphorylated MAPK-P, and so forth. The simplest assumptions about the dynamics are made. For example, take the reaction shown in the square in Figure 50. As y_3 (MAPKK-PP) facilitates the conversion of z_1 into z_2 (MAPK to MAPK-P), the rate of change $\dot{z}_1$ should include a term $-\alpha(z_1, y_3)$ (and $\dot{z}_1$

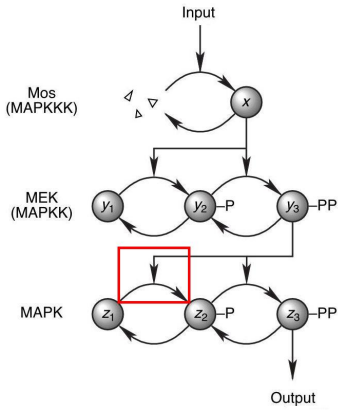

Figure 50: A MAPK system

has a term $-\alpha(z_1, y_3))$ for some (otherwise unknown) function α such that $\alpha(0, y_3) = 0$ and $\frac{\partial \alpha}{\partial z_1} > 0$, $\frac{\partial \alpha}{\partial y_3} > 0$ (more enzyme or more substrate results in a faster reaction, but nothing happens if there is no substrate). There will also be a term $+\beta(z_2)$ to reflect the phosphatase action.

As we argued in an earlier discussion, the conservation laws, here $y_1(t)+y_2(t)+y_3(t) \equiv y_{\text{tot}}$ (total MAPKK) and $z_1(t) + z_2(t) + z_3(t) \equiv z_{\text{tot}}$ (total MAPK), allow us to eliminate variables. The trick is to eliminate y_2 and z_2. Once we do this, and write $y_2 = y_{\text{tot}} - y_1 - y_3$ and $z_2 = z_{\text{tot}} - z_1 - z_3$, we are left with the variables x, y_1, y_3, z_1, z_3. For instance, the equations for z_1, z_3 look like:

$$\begin{aligned} \dot{z}_1 &= -\alpha(z_1, y_3) + \beta(z_{\text{tot}} - z_1 - z_3) \\ \dot{z}_3 &= \gamma(z_{\text{tot}} - z_1 - z_3, y_3) - \delta(z_3) \end{aligned}$$

for appropriate functions $\alpha, \beta, \gamma, \delta$. The equations for the remaining variables are similar. The incidence graph, ignoring, as usual, self-loops, is shown in Figure 51. This graph is clearly consistent, showing that the (reduced)

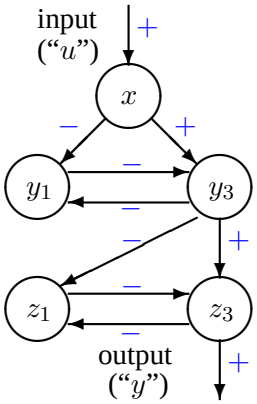

Figure 51: Graph for MAPK Reduced System

system is indeed monotone. Furthermore, the system has a well-defined characteristic, as shown in [12, 10]. Thus, the theory described here can be indeed applied.

Positive and negative feedback loops around MAPK cascades have been a topic of interest in the biological literature. For example, in [40] one finds the study of positive feedback (on Mos by ERK) in the context of progesterone-induced oocyte maturation in frogs. For another example, [60] and [81] looked at negative feedback, also in an ERK cascade, but affecting upstream proteins.

5.5.1 Numerical Example

The theorems on positive and negative feedback do not need actual equations for their applicability. All that is needed is the knowledge that the system have suitable monotonicity properties, deduced from the incidence graph, and a well-defined characteristic, plus the plot of the characteristic (steady state step response). In order to illustrate the conclusions, on the other hand, it is worth discussing a particular set of equations. Let us take the following equations:

$$\begin{aligned}
\dot{x} &= -\frac{v_2\, x}{k_2 + x} + v_0\, u + v_1 \\
\dot{y}_1 &= \frac{v_6\,(y_{\text{tot}} - y_1 - y_3)}{k_6 + (y_{\text{tot}} - y_1 - y_3)} - \frac{v_3\, x\, y_1}{k_3 + y_1} \\
\dot{y}_3 &= \frac{v_4\, x\,(y_{\text{tot}} - y_1 - y_3)}{k_4 + (y_{\text{tot}} - y_1 - y_3)} - \frac{v_5\, y_3}{k_5 + y_3} \\
\dot{z}_1 &= \frac{v_{10}\,(z_{\text{tot}} - z_1 - z_3)}{k_{10} + (z_{\text{tot}} - z_1 - z_3)} - \frac{v_7\, y_3\, z_1}{k_7 + z_1} \\
\dot{z}_3 &= \frac{v_8\, y_3\,(z_{\text{tot}} - z_1 - z_3)}{k_8 + (z_{\text{tot}} - z_1 - z_3)} - \frac{v_9\, z_3}{k_9 + z_3}
\end{aligned}$$

with output z_3. Specifically, we will use the following parameters: $v_0 = 0.0015$, $v_1 = 0.09$, $v_2 = 1.2$, $v_3 = 0.064$, $v_4 = 0.064$, $v_5 = 5$, $v_6 = 5$, $v_7 = 0.06$, $v_8 = 0.06$, $v_9 = 5$, $v_{10} = 5$, $y_{\text{tot}} = 1200$, $z_{\text{tot}} = 300$, $k_2 = 200$, $k_3 = 1200$, $k_4 = 1200$, $k_5 = 1200$, $k_6 = 1200$, $k_7 = 300$, $k_8 = 300$, $k_9 = 300$, $k_{10} = 300$. (The units are: totals in nM (mol/cm^3), v's in nM·sec^{-1} and sec^{-1}, and k's in nM.)

With these choices, *the steady state step response is the one shown in Figure 27, where y is the output z_3. Therefore, conclusions about the behavior of the closed-loop system under positive and negative feedback can be obtained by inspection of that figure.*

Figure 27 tells us that when the feedback is $u = gy$ with $g = 1/0.98$ (line of slope 0.98 when plotting y against u), there should be a unique stable state, with a high value of the output $y = z_3$, and trajectories should converge to it. Similarly, for $g = 1/2.1$ (line of slope 2.1) there should be two stable states, one with high and one with low $y = z_3$, with trajectories generically converging to one of these two. Finally, for $g = 1/6$ (line of slope 6), only the low-y stable state should persist. We verify these conclusions with four simulations in each case, for initial states (x, y_1, y_2, z_1, z_2) equal to $(170, 0, 1200, 0, 300)$ (green), $(34, 120, 600, 30, 150)$ (red), $(17, 1080, 120, 270, 30)$ (magenta), and $(0, 1200, 0, 300, 0)$ (blue). Figure 52 shows respectively the plots

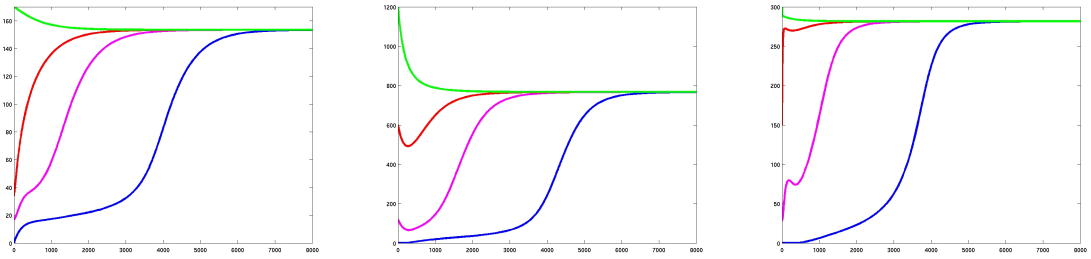

Figure 52: MAPK with positive feedback gain $g = 1/0.98$

of $x(t)$, $y_3(t)$, and $z_3(t)$ for $g = 1/0.98$, Figure 53 for $g = 1/2.1$, and Figure 54 for $g = 1/6$, confirming the theory predictions. According to the theory, the same stable states should result even if there are arbitrary delays in the feedback loop. To test this, we used the feedback law $u(t) = gz_3(t - 1000)$. We show only the simulation in the high-y case, in Figure 55.

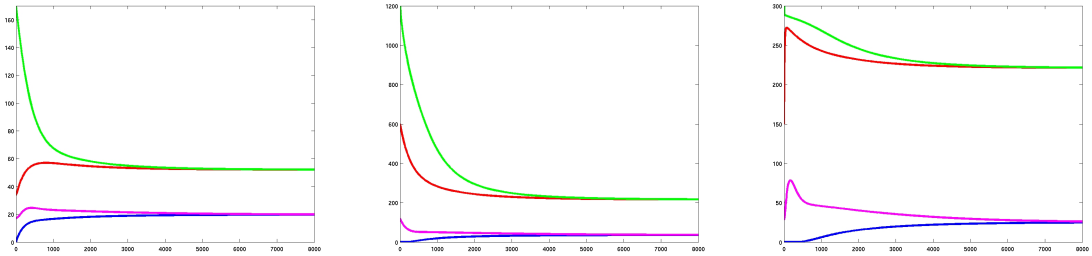

Figure 53: MAPK with positive feedback gain $g = 1/2.1$

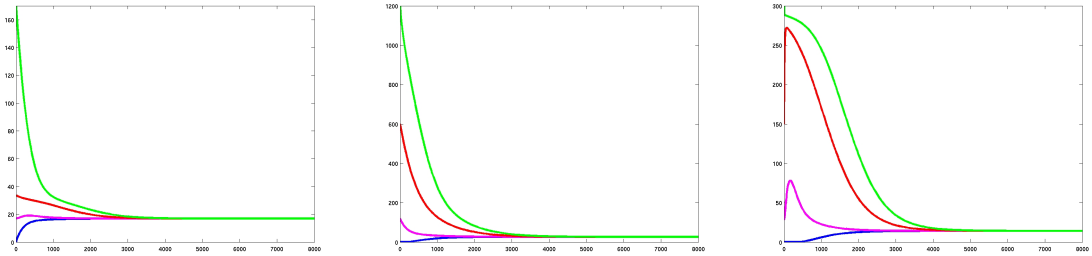

Figure 54: MAPK with positive feedback gain $g = 1/6$

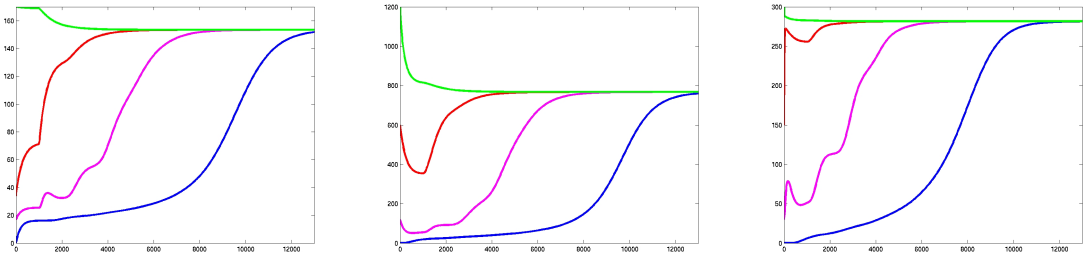

Figure 55: MAPK with delayed feedback

Next, we test the effect of diffusion. Taking, for example, the intermediate value $g = 1/2.1$, which leads to bistability, we consider the associated reaction-diffusion equation on a region as shown in Figure 56, with

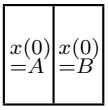

Figure 56: Initial conditions for RDE

uniform initial conditions $A = (0, 1200, 0, 300, 0)$ and $B = (170, 0, 1200, 0, 300)$ in each half that are respectively in the domains of attraction of the two predicted stable states of the closed-loop ODE. According to the theory, since the ODE solutions should converge to states with $z_3 \approx 25$ and 221, respectively, the RDE should converge to a uniform steady state equal to one of them. We plot in Figure 57 several snapshots of the time-evolution of the solution. The vertical bar shows the color coding for intensities. (The diffusion coefficient was

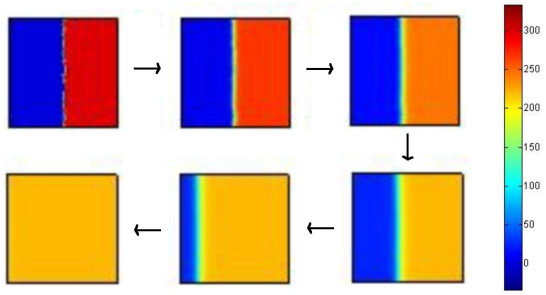

Figure 57: Reaction-diffusion solution

set to $6 \cdot 10^{-6}$ cm^2/min = 10^{-7} cm^2/sec, and the size of the square is 4cm by 4cm.) Note that each side first approaches its closest steady state, and eventually one of these (the one with $z_3 \approx 221$) takes over, confirming the theoretical prediction.

Next, we consider a *negative* feedback, namely $u = -0.9 + 600/(0.01 + z_3)$. Plotting against the characteristic, we see, through a spider-web convergence diagram for the associated discrete iteration, that the hypotheses of the negative-feedback theorem are verified, see Figure 58. Thus, the theory predicts a globally asymptoti-

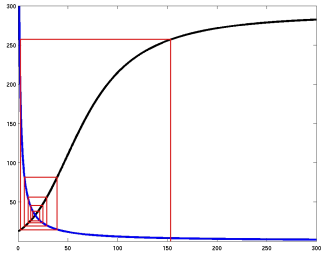

Figure 58: Discrete iteration for negative feedback

cally stable closed-loop system (even under delays and diffusion). Figure 59 shows the result of a simulation

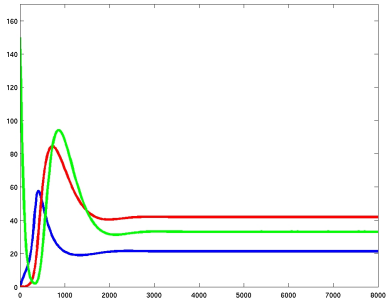

Figure 59: Negative feedback simulation

from the initial state $(0, 1200, 0, 150, 150)$, displaying the x, y_3 and z_3 coordinates together. Note that they all, indeed, converge.

Finally, we note that the bifurcation diagram for positive feedback, shown in the right panel of Figure 27, suggests that a slow adaptation law for the gain, as a negative function of z_3, should result in relaxation oscilla-

tions. This is not easy to prove, since we are dealing with a high dimensional system, but turns out to be true: a general theorem applies to guarantee the existence of such oscillations, see [43]. We do not state the precise result here, but show a simulation. We pick the adaptation law $\dot{g} = -\varepsilon(z_3 - 150)g$, with $\varepsilon = 0.000005$, and the same initial conditions as earlier. Figure 60 shows the solution trajectories for x, y_3, z_3, and Figure 61 shows

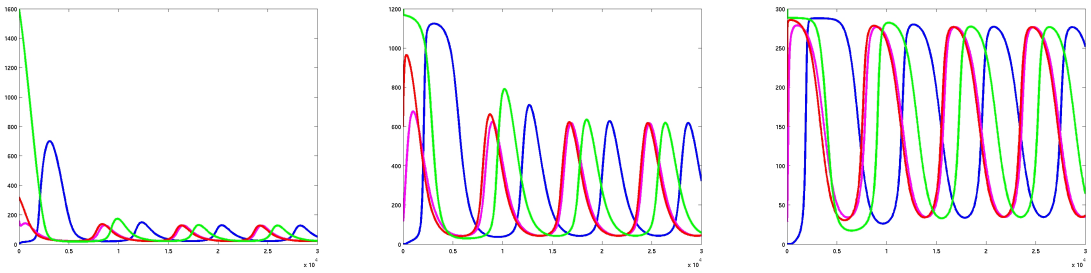

Figure 60: Relaxation oscillation, coordinates

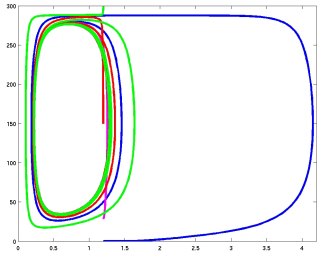

Figure 61: Relaxation oscillation, z_3, g-plane

the limit cycle behavior in the projected z_3, g plane.

6 Summary

We presented an introduction to general concepts in molecular systems biology and discussed a number of appealing dynamics and system-theoretic questions. The rapidly developing field is tremendously exciting, and full of opportunities and challenges. The reader will, hopefully, take on some of the latter.

6.1 Acknowledgments

The author wishes to thank several people who provided very useful comments and suggestions, including especially Sarah Koskie and Laura Sontag.

References

[1] Albert, R., Othmer, H.G., "The topology of the regulatory interactions predicts the expression pattern of the *Drosophila* segment polarity genes," *J. Theor. Biol.* **223**(2003), pp. 1–18.

[2] Alberts, B, Johnson, A, Lewis, J, Raff, M, Roberts, K, Walter P, *Molecular Biology of the Cell, Fourth Edition*, Garland Publishing, NY, 2002.

[3] Allwright, D.J., "A global stability criterion for simple control loops," *J. Math. Biol.* **4**(1977), pp. 363–373.

[4] Alon, U., Surette, M., Barkai, N., Leibler, S., "Robustness in Bacterial Chemotaxis," *Nature* **397**(1999), pp. 168–171.

[5] Andrec, M., B.N. Kholodenko, R.M. Levy, E. Sontag, "Inference of signaling and gene regulatory networks by steady-state perturbation experiments: Structure and accuracy," *J. Theor. Biol.* **232**(2005), pp. 427–441.

[6] Angeli, D., De Leenheer, P., Sontag, E.D., "A small-gain theorem for almost global convergence of monotone systems," *Systems and Control Letters* **51**(2004), pp. 185–202.

[7] Angeli, D., De Leenheer, P., Sontag, E.D., "Remaks on monotonicity and convergence in chemical reaction networks," in *Proc. IEEE Conf. Decision and Control, Paradise Island, Bahamas, Dec. 2004*, IEEE Publications, 2004, pp. 243–248.

[8] Angeli, D., Ferrell Jr., J.E., Sontag, E.D., "Detection of multi-stability, bifurcations, and hysteresis in a large class of biological positive-feedback systems," *Proc. Nat. Acad. Sci. USA*, **101**(2004), pp. 1822–1827.

[9] Angeli, D., Sontag, E.D., "An analysis of a circadian model using the small-gain approach to monotone systems," in *Proc. IEEE Conf. Decision and Control, Paradise Island, Bahamas, Dec. 2004*, IEEE Publications, 2004, pp. 575–578.

[10] Angeli, D., Sontag, E.D., "Interconnections of monotone systems with steady-state characteristics," in *Optimal Control, Stabilization, and Nonsmooth Analysis*, de Queiroz, M., M. Malisoff, and P. Wolenski (Eds.), (Springer-Verlag, Heidelberg, 2004), pp. 135–154.

[11] Angeli, D., Sontag, E.D., "Multistability in monotone I/O systems," *Systems and Control Letters* **51**(2004), pp. 185–202.

[12] Angeli, D., Sontag, E.D., "Monotone control systems," *IEEE Trans. Autom. Control* **48**(2003), pp. 1684–1698.

[13] Arcak, M., Sontag, E.D., "Diagonal stability for a class of cyclic systems and applications," submitted.

[14] Bagowski, C.P., Besser, J., Frey, C.R., Ferrell Jr., J.E., "The JNK Cascade as a biochemical switch in mammalian cells: ultrasensitive and all-or-none responses," *Curr. Biol.* **13**(2003), pp. 315–320.

[15] Berg, H.C., "Motile behavior of bacteria," *Physics Today*, http://www.aip.org/pt/jan00/berg.htm, January 2000.

[16] Berg, O.G., "A model for the statistical fluctuations of protein numbers in a microbial population," *J. Theor. Biol.* **71**(1978), pp. 587–603.

[17] Berman, P., Dasgupta, B., Sontag, E.D., "Randomized approximation algorithms for set multicover problems with applications to reverse engineering of protein and gene networks," *Discrete Applied Mathematics Special Series on Computational Molecular Biology* (in press).

[18] Bhalla, U.S., Ram, P.T., Iyengar, R., "MAP kinase phosphatase as a locus of flexibility in a mitogen-activated protein kinase signaling network," *Science* **297**(2002), pp. 1018–1023.

[19] http://www.cellsignal.com.

[20] Chaves, M., Albert, R., Sontag, E.D., "Robustness and fragility of Boolean models for genetic regulatory networks," *J. Theoretical Biology* **235**(2005), pp. 431–449.

[21] Chaves, M., Dinerstein, R.J., Sontag, E.D., "Optimal length and signal amplification in weakly activated signal transduction cascades," *J. Physical Chemistry* **108**(2004), pp. 15311–15320.

[22] Chaves, M., Sontag, E.D., "State-Estimators for chemical reaction networks of Feinberg-Horn-Jackson zero deficiency type," *European J. Control* **8**(2002), pp. 343–359.

[23] Chaves, M., Sontag, E.D., "Computation of amplification for systems arising from cellular signaling pathways," in *Proc. 16th IFAC World Congress*, (Prague, July 2005), to appear. (See also: arxiv.org q-bio.QM/0408012 Aug 2004)

[24] Chaves, M., Sontag, E.D., Dinerstein, R.J., "Gains and optimal design in signaling pathways," in *Proc. IEEE Conf. Decision and Control, Paradise Island, Bahamas, Dec. 2004*, IEEE Publications, 2004, pp. 596–601.

[25] Chaves, M., Sontag, E.D., Dinerstein, R.J., "Steady-states of receptor-ligand dynamics: A theoretical framework," *J. Theoretical Biology* **227**(2004), pp. 413–428.

[26] Cinquin, O., Demongeot, J., "Positive and negative feedback: striking a balance between necessary antagonists," *J. Theor. Biol.* **216**(2002), pp. 229–241.

[27] Dancer, E.N., "Some remarks on a boundedness assumption for monotone dynamical systems," *Proc. of the AMS* **126**(1998), pp. 801–807.

[28] B. DasGupta, G.A. Enciso, E.D. Sontag, Y. Zhang, "On approximate consistent labeling of biological dynamical systems," submitted.

[29] Davidson, E.H., et al., "A genomic regulatory network for development," *Science* **295**(2002), pp. 1669–1678.

[30] De Leenheer, P., Angeli, D. Sontag, E.D., "Monotone chemical reaction networks," *J. Mathematical Chemistry*, to appear.

[31] De Leenheer, P., Angeli, D., Sontag, E.D., "A tutorial on monotone systems- with an application to chemical reaction networks," in *Proc. 16th Int. Symp. Mathematical Theory of Networks and Systems (MTNS 2004)*, CD-ROM, WP9.1, Katholieke Universiteit Leuven.

[32] De Leenheer, P., Angeli, D. Sontag, E.D., "On predator-prey systems and small-gain theorems," *J. Mathematical Biosciences and Engineering* **2**(2005), pp. 25–42.

[33] Devreotes, P., Sherring, J., "Kinetics and concentration dependence of reversible cAMP-induced modification of the surface cAMP receptor in *Dictyostelium*," *J. Biol. Chem.* **260**(1985), pp. 6378–6384.

[34] Edelstein-Keshet, L. *Mathematical Models in Biology*, McGraw-Hill, 1988

[35] Enciso, G.A., Smith, H.L., Sontag, E.D., "Non-monotone systems decomposable into monotone systems with negative feedback," *J. Differential Equations*, to appear. (doi:10.1016/j.jde.2005.05.007)

[36] Enciso, G.A., Sontag, E.D., "Global attractivity, I/O monotone small-gain theorems, and biological delay systems," *Discrete and Continuous Dynamical Systems* **14**(2006), pp. 549–578.

[37] Enciso, G.A., Sontag, E.D., "Monotone systems under positive feedback: multistability and a reduction theorem," *Systems and Control Letters* **54**(2005), pp. 159–168.

[38] Enciso, G.A., Sontag, E.D., "On the stability of a model of testosterone dynamics," *J. Mathematical Biology* **49**(2004), pp. 627–634.

[39] Feinberg, M., "Chemical reaction network structure and the stabiliy of complex isothermal reactors - I. The deficiency zero and deficiency one theorems," *Chemical Engr. Sci.* **42**(1987), pp. 2229–2268.

[40] Ferrell Jr., J.E., Machleder, E.M., "The biochemical basis of an all-or-none cell fate switch in Xenopus oocytes," *Science* **280**(1998), pp. 895–898.

[41] Ferrell Jr., J.E., "Tripping the switch fantastic: How a protein kinase cascade can convert graded inputs into switch-like outputs," *Trends Biochem. Sci.* **21**(1996), pp. 460–466.

[42] Gedeon, T., "Cyclic Feedback Systems," *Mem. Amer. Math. Soc.* **13**(1998), pp. 1–73.

[43] Gedeon,T., Sontag, E.D., "Oscillation in multi-stable monotone system with slowly varying positive feedback," in preparation. (Abstract in *Sixth SIAM Conf. on Control and its Applications*, New Orleans, July 2005.)

[44] Goldbeter, A., *Biochemical Oscillations and Cellular Rhythms*, Cambridge University Press, Cambridge, 1996.

[45] Gouzé, J.L.. "Positive and negative circuits in dynamical systems," *J. Biological Systems* **6**(1998), pp. 11–15.

[46] Hanahan, D., Weinberg, R.A., "The hallmarks of cancer," *Cell* **100**(2000), pp. 57–70.

[47] Hartwell, L.H., Hopfield, J.J., Leibler, S., Murray, A.W., "From molecular to modular cell biology," *Nature* **402**(1999), pp. C47–C52.

[48] Hastings, S., Tyson, J., Webester, D., "Existence of periodic solutions for negative feedback cellular control systems," *J. Diff. Eqs.* **25**(1977), pp. 39–64.

[49] Heinrich, R., Neel, B.G., Rapoport, T.A., "Mathematical models of protein kinase signal transduction," *Molecular Cell* **9**(2002), pp. 957–970.

[50] Hirsch, M., "Systems of differential equations that are competitive or cooperative II: Convergence almost everywhere," *SIAM J. Mathematical Analysis* **16**(1985), pp. 423–439.

[51] Hirsch, M., "Differential equations and convergence almost everywhere in strongly monotone flows," *Contemporary Mathematics* **17**(1983), pp. 267–285.

[52] Hirsch, M., H.L. Smith, "Monotone dynamical systems," in *Handbook of Differential Equations, Ordinary Differential Equations (second volume)*, Elsevier, 2005, to appear.

[53] Horn, F.J.M., Jackson, R., "General mass action kinetics," *Arch. Rational Mech. Anal.* **49**(1972), pp. 81–116.

[54] Huang, C.-Y.F., Ferrell Jr., J.E., "Ultrasensitivity in the mitogen-activated protein kinase cascade," *Proc. Natl. Acad. Sci. USA* **93**(1996), pp. 10078–10083.

[55] Iglesias, P.A., Levchenko, A., "A general framework for achieving integral control in chemotactic biological signaling mechanisms," Proc. IEEE Conf. Decision and Control, 2001, Orlando, USA, pp. 843–848.

[56] Jiang, J.F., "On the global stability of cooperative systems," *Bulletin of the London Math Soc* **6**(1994), pp. 455–458.

[57] Kauffman, S.A., "A proposal for using the ensemble approach to understand genetic regulatory networks," *J. Theoretical Biololy* **230**(2004), pp. 581–590

[58] Keener, J.P., J. Sneyd, *Mathematical Physiology*, Springer-Verlag New York, 1998.

[59] Kepler, T.B., Elston, T.C., "Stochasticity in transcriptional regulation: origins, consequences, and mathematical representations," *Biophys J.* **81**(2001), pp. 3116–3136.

[60] Kholodenko, B.N., "Negative feedback and ultrasensitivity can bring about oscillations in the mitogen-activated protein kinase cascades," *Eur. J. Biochem* **267**(2000), pp. 1583–1588.

[61] Kholodenko, B.N., Kiyatkin, A., Bruggeman, F., Sontag E.D., Westerhoff, H., Hoek, J., "Untangling the wires: a novel strategy to trace functional interactions in signaling and gene networks," *Proceedings of the National Academy of Sciences USA*, **99**(2002), pp. 12841-12846.

[62] Lauffenburger, D.A., "A computational study of feedback effects on signal dynamics in a mitogen-activated protein kinase (MAPK) pathway model," *Biotechnol. Prog.* **17**(2001), pp. 227–239.

[63] Lauffenburger, D.A., "Cell signaling pathways as control modules: Complexity for simplicity," *Proc. Nat. Acad. Sci. USA* **97**(2000), pp. 5031-5033.

[64] Maayan, A., Iyengar, R. Sontag, E.D., "Sign-consistency loops detected in biochemical regulatory networks may provide dynamical stability," in preparation.

[65] Mallet-Paret, J., Smith, H.L., "The Poincaré-Bendixson theorem for monotone cyclic feedback systems," *J. Dynamics and Diff. Eqns.* **2**(1990), pp. 367–421.

[66] Mangan, S., Alon, U., "Structure and function of the feed-forward loop network motif," *Proc. Natl. Acad. Sci. (USA)* **100**(2003), pp. 11980–11985.

[67] Mangan, S., Zaslaver, A., Alon, U., "The coherent feedforward loop serves as a sign-sensitive delay element in transcription networks," *J. Molec. Bio.* **334**(2003), pp. 197–204.

[68] McKeithan, T.W., "Kinetic proofreading in T-cell receptor signal transduction," *Proc. Natl. Acad. Sci. USA* **92**(1995), pp. 5042-5046.

[69] Milo, R., Shen-Orr, S., Itzkovitz, S., Kashtan, N., Chklovskii, D., Alon, U., "Network motifs: Simple building blocks of complex networks," *Science* **298**(2002), pp. 824–827.

[70] Murray, J.D., *Mathematical Biology, 3rd Edition*, Springer, New York, 2002.

[71] Novick, A., Wiener, M., "Enzyme induction as an all-or-none phenomenon," *Proc. Natl. Acad. Sci. U.S.A.* **43**(1957), pp. 553–566.

[72] Othmer, H.G, "The qualitative dynamics of a class of biochemical control circuits," *J. Math. Biol.* **3**(1976), pp. 53–78.

[73] Pomerening, J.R., Sontag, E.D., Ferrell Jr., J.E., "Building a cell cycle oscillator: hysteresis and bistability in the activation of Cdc2," *Nature Cell Biology* **5**(2003), pp. 346–351.

[74] Ptashne, M., *A Genetic Switch: Phage λ and Higher Organisms*, (Cell Press and Blackwell Scientific Publications, Cambridge MA), 1992.

[75] Rao, C.V., Wolf, D.M., Arkin, A.P., "Control, exploitation and tolerance of intracellular noise," *Nature* **420**(2002), pp. 231–237.

[76] Rapp, P.E., "A theoretical investigation of a large class of biochemical oscillations," *Math Biosciences* **25**(1975), pp. 165–188.

[77] Saez-Rodriguez, J., Kremling, A., Conzelmann, H., Bettenbrock, K., Gilles, E.D., "Modular analysis of signal transduction networks," *IEEE Control Systems Magazine* **24**(2004), pp. 35–52.

[78] Segall, J.E., Block, S.M., Berg, H.C., "Temporal comparisons in bacterial chemotaxis," *Proc. Nat. Acad. Sci. USA* **83**(1986), pp. 8987–8991.

[79] Seger, R., "Intracellular signaling cascades,," http://www.weizmann.ac.il/Biology/open_day_2002/book/rony_seger.pdf.

[80] Sha, W., Moore, J., Chen, K., Lassaletta, A.D., Yi, C.S., Tyson, J.J., Sible, J. C., "Hysteresis drives cell-cycle transitions in Xenopus laevis egg extracts," *Proc. Natl. Acad. Sci. USA* **100**(2003), pp. 975–980.

[81] Shvartsman, S.Y., Wiley, H.S., Lauffenburger, D.A., "Autocrine loop as a module for bidirectional and context-dependent cell signaling," preprint, MIT Chemical Engineering Department, December 2000.

[82] Smith, H.L., "Oscillations and multiple steady states in a cyclic gene model with repression," *J. Math. Biol.* **25**(1987), pp. 169–190.

[83] Smith, H.L., *Monotone dynamical systems: An introduction to the theory of competitive and cooperative systems, Mathematical Surveys and Monographs, vol. 41*, (AMS, Providence, RI, 1995).

[84] Shmulevich, I., Lähdesmäki, H., Dougherty, E.R., Astola, J., Zhang, W., "The role of certain Post classes in Boolean network models of genetic networks," *Proc. National Acad. Sci. USA* **100**(2003), pp. 10734–10739

[85] Shmulevich, I., Kauffman, S.A., "Activities and sensitivities in Boolean network models," *Phys. Rev. Lett.* **93**(2004), pp. 048701–4

[86] Snoussi, E.H., "Necessary conditions for multistationarity and stable periodicity," *J. Biol. Sys.* **6**(1998), pp. 3–9.

[87] Sontag, E.D., "Spaces of observables in nonlinear control," in *Proc. Intern. Congress of Mathematicians 1994*, Volume 2, Birkhäuser Verlag, Basel, 1995, pp. 1532-1545.

[88] Sontag, E.D., *Mathematical Control Theory: Deterministic Finite Dimensional Systems*, Springer, New York, 1990. Second Edition, 1998.

[89] Sontag, E.D., "Structure and stability of certain chemical networks and applications to the kinetic proof-reading model of T-cell receptor signal transduction," *IEEE Trans. Autom. Control* **46**(2001), 1028–1047. Errata in *IEEE Trans. Autom. Control* **47**(2002), pp. 705–705.

[90] Sontag, E.D., "For differential equations with r parameters, $2r+1$ experiments are enough for identification," *J. Nonlinear Science* **12**(2002), pp. 553–583.

[91] Sontag, E.D., "Adaptation and regulation with signal detection implies internal model," *Systems and Control Letters* **50**(2003), pp. 119–126.

[92] Sontag, E.D. "Some new directions in control theory inspired by systems biology," *Systems Biology* **1**(2004), pp. 9–18.

[93] Sontag, E.D., Kiyatkin, A., Kholodenko, B.N., "Inferring dynamic architecture of cellular networks using time series of gene expression, protein and metabolite data," Bioinformatics **20**(2004), pp. 1877–1886

[94] Stock, A.M., V.L. Robinson, P.N. Goudreau, "Two-component signal transduction," *Annual Review Biochem.* **69**(2000), pp. 183–215.

[95] Thattai, M., van Oudenaarden, A., "Intrinsic noise in gene regulatory networks," *Proc. Natl Acad. Sci. USA* **98**(2001), pp. 8614-8619.

[96] Thomas, R., in *Springer Series in Synergetics 9*, Springer-Verlag, Berlin, 1981, pp. 180–193.

[97] Thron, C.D., "The secant condition for instability in biochemical feedback-control. 1.The role of cooperativity and saturability," *Bull. Math. Biology* **53**(1991), pp. 383–401.

[98] Tyson, J.J., Othmer, H.G., "The dynamics of feedback control circuits in biochemical pathways," in *Progress in Theoretical Biology Vol. 5 (R. Rosen & F.M. Snell, eds.),*, Academic Press, New York, 1978, pp. 1–62.

[99] van Riel, N.A.W., Sontag, E.D., "Parameter estimation in dynamic models combining signal transduction and metabolic pathways: The dependent input approach," submitted.

[100] Volpert, A.I., Volpert, V.A., Volpert, V.A., *Traveling Wave Solutions of Parabolic Systems*, Mathematical Monographs, Vol. 140, AMS, Providence (RI).

[101] von Dassow, G., Meir., E., Munro, E.M., Odell, G.M., "The segment polarity network is a robust developmental module," *Nature* **406**((2000), pp. 188–192.

[102] Widmann, C., Spencer, G., Jarpe, M.B., Johnson, G.L., "Mitogen-activated protein kinase: Conservation of a three-kinase module from yeast to human," *Physiol. Rev.* **79**(1999), pp. 143–180.

[103] Wolpert, L., Beddington, R., Brockes, J., Jessell, T., Lawrence, P., Meyerowitz, E., *Principles of Development*, Current Biology Ltd., London, 1998.

[104] Yi, T.-M., Huang, Y., Simon, M.I., Doyle, J., "Robust perfect adaptation in bacterial chemotaxis through integral feedback control," *Proc. Natl. Acad. Sci. USA* **97**(2000), pp. 4649–4653.

[105] Yuh, C.H., Bolouri, H., Bower, J.M., Davidson, E.H., "A logical model of cis-regulatory control in a eukaryotic system," in *Computational Modeling of Genetic and Biochemical Networks*, (eds. Bower, J.M. and Bolouri, H.) MIT Press, Cambridge, MA, 2001, pp. 73–100.